

ISSN 2510-2591



Reports of the European Society for
Socially Embedded Technologies

volume 6 issue 1
2022

Proceedings of 20th European Conference on Computer-Supported Cooperative Work

Guest Editors

Paula Alexandra Silva, University of Coimbra, Portugal
Nervo Xavier Verdezoto, Cardiff University, UK

Series Editor

Michael Koch

Impressum

The '**Reports of the European Society for Socially Embedded Technologies**' are an online report series of the European Society for Socially Embedded Technologies (EUSSET). They aim to contribute to current research discourses in the fields of 'Computer-Supported Cooperative Work', 'Human-Computer-Interaction' and 'Computers and Society'.

The 'Reports of the European Society for Socially Embedded Technologies' appear at least one time per year and are exclusively published in the Digital Library of EUSSET (<https://dl.eusset.eu/>). The main language of publication is English.

ISSN 2510-2591

<https://www.eusset.eu/report-series/>

The contents of the report series are made available in the EUSSET Digital Library at <https://dl.eusset.eu/>

EUSSET is an institute of Social Computing e.V., a non-profit association according to the German legal system – founded on November 13th 2012 in Bonn, Germany (Nordrhein-Westfalen Amtsgericht Bonn VR 9675).

c/o Prof. Dr. Volker Wulf
Fakultät III
Universität Siegen
57068 Siegen
E-Mail: volker.wulf@uni-siegen.de

Table of Contents

Exploratory Papers

I think "Hedging" could be a Feminist Issue in Software Engineering

Ashcroft, Alice

Shall I describe it or shall I move closer? Verbal references and locomotion in VR collaborative search tasks.

Bovo, Riccardo; Giunchi, Daniele; Costanza, Enrico; Steed, Anthony; Heinis, Thomas

Redesigning systems for Single-Pilot Operations: the mutual awareness problem for remote crews

Bardou, Maxime; Letondal, Catherine; Imbert, Jean-Paul; Causse, Mickaël; Bidegaimberry, Maxence; Dubus, Romane; Marcon, Cécile

Toward an AI-assisted Assessment Tool to Support Online Art Therapy Practices: A Pilot Study

Seo, Woosuk; Jun, Joonyoung; Chun, Minki; Jeong, Hyeonhak; Na, Sungmin; Cho, Woohyun; Kim, Saerj; Jung, Hyunggu

Comparing different qualitative methods to understand user experience in Saudi Arabia

AlArfaj, Aisha Ahmed; Solaima, Ellis

Travelling Artefacts: Lessons Learned from Interventions in a Regional Innovation Ecosystem

Krüger, Max; Gerbracht, Marc; Vitt, Nico; Kudic, Muhamed; Ahmadi, Michael; Boden, Alexander; Offergeld, Felicitas; Stein, Martin; Kotthaus, Christoph; Unbehaun, David; Wulf, Volker

The Role of Boundary Objects in Platformization Practices: A Case Study of Software Testing

Ringdal, Nora O.; Farshchian, Babak A.

The potential of a digital twin to design computational coordination mechanisms. The case of the French railway infrastructure

Stalder, Corentin; Lewkowicz, Myriam; Ducellier, Guillaume

Connectedness in mobile families

Glöss, Mareike

Causal Impact Model to Evaluate the Diffusion Effect of Social Media Campaigns

Yu, Xinchun; Mashhadi, Afra; Boy, Jeremy; Nielsen, Rene Clausen; Hong, Lingzi

Pictorial Consent: Fieldwork Reflections

Saha, Manika; Bartindale, Tom; Sultana, Sharifa; Oliver, Gillian; Thilsted, Shakuntala Haraksingh; Ahmed, Syed Ishtiaque; Varghese, Delvin; Olivier, Patrick

The Design Multiple: Sharedness and Multiplicity in Common Information Space

Mohallick, Itishree; Monteiro, Eric; Østerlie, Thomas

Coupling documentation and communication practices to support integrated care pathways

Marref, Rahma; Amsha, Khuloud Abou; Lewkowicz, Myriam

Notes

Enterprise Collaboration Platform Configurations: an Empirical Study

Schubert, Petra; Williams, Susan P.

Toward Supporting the Mental Health of Underprivileged Youth Through Village-Driven Sociotechnical Systems

Huh-Yoo, Jina; Montgomery, Meredith; Matta, Rishabh; Kim, Heejun; Daly, Brian

Deconstructing Gender in Asylum Categories: An Archival Perspective on a Practice with Limited Access

Kaltenhäuser, Kristin; Slaats, Tijs; Gammeltoft-Hansen, Thomas; Holten Møller, Naja

Blending Practices to Facilitate Grounded Design Research: A Praxeological Research Perspective

Syed, Hussain Abid; Schorch, Marén; Pinatti de Carvalho, Aparecido Fabiano; Rutz, Philipp; Pipek, Volkmar

Toward Trauma-Informed Design of Behavioral Interventions: A Case Study on Classroom Management

Marcu, Gabriela

Alice Ashcroft (2022): I think "Hedging" could be a Feminist Issue in Software Engineering. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_ep01

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_ep01

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

I think "Hedging" could be a Feminist Issue in Software Engineering

Alice Ashcroft

School of Computing and Communication, Lancaster University

Contact Author: alice.ashcroft@lancaster.ac.uk

Abstract. When it comes to software engineering and the development life-cycle, there are a number of opportunities for under-represented groups, gender being the focus of this paper, for decisions to be affected by language. Considering existing linguistics research surrounding gendered language, specifically "Hedging", alongside various stages of the development life-cycle, this paper poses that "Hedging" should be seen as a feminist issue in software engineering, and presents five areas for further research to uncover the potential negative effects it is having, and what can be done to mitigate these. This paper focuses on the subtleties in conversation, and how conversation takes place, building on Feminist Conversation Analysis, Feminist Methodologies, and Software Engineering Methodologies.

1 Introduction

The effect that gender, language and interaction can have on Computer Science (CS) cannot be understated. From initial meetings where ideas are brought forward, to design processes, to the systems being built, to the impact this has on users; each of these stages, and the rest, rely heavily on communication. Sometimes, these interactions are textual, frequently face to face, and more and more since the Covid-19 pandemic began, through video calls.

Language, outside of CS has long been studied, particularly when it comes to the effects of gender¹. However, when looking at the direct impact this has on the software being built, the findings of this are much more recent and have only since been recognised as holding some importance. Whilst the effect of communication on Software Engineering has begun to be explored (Catolino et al., 2019), the subsection of gendered language which has yet to be examined in much detail, and is therefore reflected upon in this exploratory paper, is that of "**Hedging**". "Hedging" is short for "*hedging your bets*" and is often displayed through the use of phrases such as, "*I think*", "*you know*" and "*sort of*" (Holmes, 1986; Tannen, 1994; Murphy, 2010; Ashcroft, 2020b, 2021). There are, of course, many other gender differences in language (as outlined in Section 3.2), but the focus of this paper will be Hedging, due to the potential impact for miscommunication it may cause.

It is also important to consider these through the lens of Feminist Methodologies (discussed further in Section 2), and how these are already applied within CS, and how an application of these can contribute to the overall knowledge within the field, and in this particular instance the Software Development Life-cycle.

Therefore, this paper will focus on the Software Development Life-cycle, also known as the Software Process, as outlined by Sommerville (2016), and discusses how discourse, interaction and language have the potential to impact decisions and processes throughout. However, in order to do this, there first needs to be an understanding of Feminism in CS Methodologies (Section 2) and following this, the impact gender can have on language (Section 3).

2 Feminism in Computer Science Methodologies

It could be argued that Feminism in Computer Science and Feminist Methodologies are simply good research practices, however, Feminist Methodologies operate under the assumption that any research used as a basis for understanding may contain unconscious patriarchal bias (O'Leary, 2017). This assumption is supported by Davis (2019), who argues that diversity is "not only essential as a goal of intersectional feminism², but also as a means of continuing success for digital movements." Therefore, in order to overcome any existing bias in research, any research carried out going forward should not rely on past work, although many researchers would agree that this is rarely possible.

¹ Throughout this paper, the emphasis is on gender, and not sex, however much existing literature does not differentiate between the two. Where possible the terms "men", "women" and "gender" have been used, instead of "male", "female" and "sex", however, the cases where they are used are done so to correctly reflect previous work. This author also rejects the idea that gender is a binary, despite the fact that most existing literature refers to it as such. Again, literature will be referenced accurately, but this should not undermine the need for more extensive research into gender beyond the binary and these issues.

² Although linguistically, intersectionality may simply mean the cross over of multiple identities, when it comes to theory, this is a product of Black Feminist Theory, and therefore this should be recognised and credited, otherwise this risks the loss of credit from already underrepresented groups (De Hertogh et al., 2019).

Furthermore, Schlesinger et al. (2017) introduce intersectionality as a framework for research, recognising it as the overlap of attributes making up an individual's identity. They carried out further research into the intersectionality of HCI research by bringing together a keyword set and "included terms that describe gender, ethnicity, race, sexuality or class" and found that only 1% of all CHI publications included these keywords, showing that work has yet to be done in this area (Schlesinger et al., 2017). This is also supported by Bellini et al. (2018) who state that "an inclusive community must be built that connects research from across different areas within HCI, across academic generations, and across other tangible and intangible borders", an idea that was previously discussed at the EUSSET Colloquium on "CSCW Theories and Concepts" in January of 2022 (eus, 2022) in that research must be carried out regarding diversity, but it must also be part of other research where gender is not the focus.

In addition to this, when it comes to Feminist Epistemology, "there has been debate between feminists about whether there can be feminist epistemology" (Barbour, 2018), but if epistemology refers to the theory of knowledge and understanding, and the subjects of the research "have gender", then the argument from feminists that "gender and individual identity are significant in the process of becoming a subject and a knower" (Flax, 1993; Barbour, 2018), then surely they must be relevant in social research? However when it comes to applying feminist epistemologies in CS, where often logic and structure are seen as important values, this may clash with the more general understanding of how knowledge is formed. As stated by Hancox-Li and Kumar (2021), "feminist epistemology has long taken a critical stance towards fully formalized systems, instead emphasizing the interactive nature of knowledge creation and the importance of exploring multiple possible meanings". This could be argued to be even more important to consider where gendered language is the topic of research, as it is in this paper, as the difference in meaning and understanding of both the language used and the research being carried out, must be held to the same standards.

Therefore, considering feminist theories, and feminist research, and how this can support further research into language and its impact on design, particularly within a field where logic is valued so highly, a closer examination of literature regarding CA and DA has been carried out (Section 3), in order to provide a framework for uncovering how "Hedging" can have an impact on design in CS (as discussed in Section 4).

3 Gender, Language and Interaction

Conversation can happen in a multitude of ways, but there are classic traits of conversation that have been observed and researched through both Conversation Analysis (CA) and Discourse Analysis (DA), as discussed in Section 3.1. These traits, like so many things, are affected by gender (see Section 3.2), and one of these traits is known as "**Hedging**" (see Section 3.3). This Section will explore the classic traits of conversation from the field of linguistics, how gender affects these,

and how "Hedging" is an example of how this can have an impact on the decisions made throughout conversation. Following this examination of literature regarding gendered language, and hedging, this understanding will then be taken into Section 4 and analysed alongside the Software Development Cycle.

3.1 Traits of Conversation

Traits of conversation, and discourse, as stated above, are well established outside of the field of CS. Stokoe (2018) is just one of many to look at these interactions and traits, particularly when it comes to gender, with others including Holmes (1986) and Boden (1994). Opening conversation, turn takings, gaps and overlaps as well as how the conversation ends, are how conversations take place (Stokoe, 2018) and the impact that gender can have on these parts of conversation can be affected by sex differences in language, gender itself, how problems may be raised, and "**Hedging**".

How a **conversation opens** can have a significant impact on how the rest of the conversation is likely to go (Stokoe, 2018). There is seemingly little research on how gender can affect this well-known phenomenon discussed often in CA (Svennevig, 2012). **Turn-taking**, however, can largely be affected by gender due to the effect power dynamics can play on this trait; which could be caused by any number of attributes including; experience in a role, race, gender, etc. Ford (2008) found that in academic meetings the majority of turn-taking was decided through non-verbal cues "such as leaning forward, gazing at the chair and raising a hand" (Svennevig, 2012). Turn-taking, of course, may be impacted by a large number of meetings now taking place online. It could be argued that having meetings online may allow for more equal participation, due to these physical barriers being removed, or perhaps the more outspoken members of meetings are more likely to speak and therefore be heard, due to the lack of, or perhaps increased, opportunity for overlap in conversation. Paulus et al. (2016) analysed literature relating to online conversation analysis before the pandemic in 2016 and relate this back to the work of Sacks et al. (1974) and whilst finding some deviations, find a number of similarities in how turn taking takes place online compared to in person.

On the topic of **gaps and overlaps** in conversation, both can have a serious effect on the conclusions reached in a conversation (Liddicoat, 2011), especially when there are more than two people in the conversation. Overlaps in conversation could be seen to have rather negative connotations, anecdotally it seems that nobody enjoys being spoken over, but Stokoe (2018) argues that overlap can be "an example of collaboration", which could be argued as being vital to any software development process.

How **conversation closes** is often the last step of any conversation (Liddicoat, 2011; Jay, 2009), and when it comes to meetings, is often concluded by a summary of actions or leanings. Therefore it is vital to have a "good conversation" or "good meeting", where everybody leaves with the same understanding of what the

conversation has achieved; be it the exchanging of pleasantries or the recording of actions.

3.2 Gender and Language

Gender and language is an established field, with many researchers using CA and DA to uncover gender differences (Stokoe and Smithson, 2001; Benwell, 2006; Friedrich and Heise, 2019). Research has also started to be done examining the effect that gender and language can have on User Experience (UX) and Software Design (Ashcroft, 2020a). Therefore, it is vital to consider the effect gender has on the conversations that take place when gender is not the topic of conversation.

Speer and Stokoe (2011) state that one consideration of gender and language is "**sex differences in language**", which focuses on the way men and women are represented in the language itself. For example, referring to a generic user as "*he*", as opposed to using gender-neutral pronouns such as "*they*". **How gender, and varying identities, are constructed** is another area that affects discourse (Speer and Stokoe, 2011).

How problems are raised will have an impact on how groups make decisions (Park, 1996), and this has also been linked to gender (Ashcroft, 2020a). This could also have a major effect on software development processes, if women in the meeting feel unable to raise problems, or if they do they are not heard.

Following from the traits of conversation discussed above (Section 3.1), how people **overlap** in conversation and **take turns**, can also be impacted by gender (Kitzinger, 2008). For example, the order of speakers may seemingly be chosen at random (e.g. going around a table), but this may sometimes be led by the men who are present (Ashcroft, 2020a). Therefore, it should be considered who is leading the discussion, which may again tie into the roles of power, and existing dynamics caused by roles, gender, race etc. - especially given that "men often achieve leadership roles regardless of past performance" (Reuben et al., 2012).

This paper poses that each of these existing areas for potential inequality will only be worsened by the effects that "Hedging" may have on conversation³.

3.3 What is "Hedging"?

As outlined above, "Hedging" is short for "hedging your bets" and includes the use of phrases such as, "I think", "you know" and "sort of" (Holmes, 1986; Tannen, 1994; Murphy, 2010; Ashcroft, 2020b, 2021). Murphy (2010) states that "the use of hedges among females before a key word" is used "to avoid the appearance of playing the expert", also known as "Expressing Uncertainty" (Ashcroft, 2020b). Nearly all researchers agree that "Hedging" will be used with different motivations

³ It could be argued, that due to the existing inequality in the perception of talk, that "Hedging" could be used as a tool to encourage ideas and thoughts to be heard; for example being used to persuade, or to make the words of those who are under-represented more "palatable" for others. Therefore, "Hedging" should be seen as a tool for conversation, not always a hindrance, which is discussed further in Section 3.3

and that context plays a part in understanding the reason for its use (Holmes, 1990). For example, Holmes (1986) states that "Hedging" can have two purposes, either to express speaker confidence e.g. "you know" or "reflecting uncertainty" e.g. "I think". Regardless of the motivation, "Hedging" is recognised as a "significant communicative resource for academics" (Hyland, 1996), understanding "Hedging", the motivations for its use and the impact it can have on conversation and decisions, is vital to understanding how decisions are made in the real world, and in the case of this paper, in design.

"Hedging" can have an impact on discussions that take place, because it is mainly rooted in the aim of downplaying ideas, thoughts or suggestions. For example in past research, focusing on design meetings, a woman said "some lecturers don't use the interactive screen, so, um, so *I think* you want more people to use it..." (Ashcroft, 2020b). This example could be seen as the expression of genuine uncertainty, or they could have been attempting to reduce their contribution for fear of rejection or taking up space, which is a known feminist issue, one that affects Black Women even more (Blackwell, 2020).

The impact this may have cannot be understated, not only will this have an impact on design decisions that are made, but could potentially disrupt the entire process and negate or downplay the opinions portrayed by women. However, it is important that there is no assumption that "Hedging" is itself negative, as well as being used to express uncertainty, it can also be used to persuade, with phrases such as "you know". Therefore, the topic that is focused on in this paper (see Section 4) is the perception of the words being spoken, regardless of the they *type* of "Hedging" being used, particularly when it comes to gender differences.

Dixon and Foster (1997) found that "Hedging" was used the same number of times by both genders but the reason *why* it was used differed based on gender, which contradicts what was found by Holmes (1986), who found that "Hedging" was used more by women. It could be argued that the frequency of use is irrelevant without understanding the reasons why it was used. However, there seems to be little research in which participants have been interviewed after a recording has taken place, in which their motivations are questioned, and all analysis seems to be inferred by the researcher's digression (Holmes, 1986; Ashcroft, 2020b). One of the main issues with this research approach, however, could be that participants are not sure of their reasons for saying it, as it may be instinctive or done without thinking. Therefore, for intent to be uncovered, this will require further research.

All this being said, it is clear that there is a correlation in existing research between gender, "Hedging", and the effect this has on a conversation. It is these parallels that will be discussed alongside the software development cycle in Section 4, in order to uncover its potential impact. All of this, of course, should then be tested through the recording and analysis of real design meetings, to ensure its validity, as discussed in Section 6.

4 The Potential Effects of "Hedging" on the Software Development Life-cycle

The Software Development Life-cycle (SDL) is a well-documented process (Sommerville, 2016; Leau et al., 2012), and whilst variations on this exist such as through Innovation Practices (Kic, 2021), Agile Methodologies (Agile Alliance, 2020) and the more traditional Waterfall (Dima and Maassen, 2018); the fundamental principles supporting this remain the same. These are as follows;

1. "**Software Specification:** The functionality of the software and constraints on its operation must be defined." (Sommerville, 2016)
2. "**Software Development:** The software to meet the specification must be produced." (Sommerville, 2016)
3. "**Software Validation:** The software must be validated to ensure that it does what the customer wants." (Sommerville, 2016)
4. "**Software Evolution:** The software must evolve to meet changing customer needs." (Sommerville, 2016)

There, of course, exists the process before Specification even begins, where ideas or problems are brought forward to begin with (see Section 4.1). It could be argued that this is the true start of the "pipeline". Defined as Ideation by Adobe (Kic, 2021), problem raising through other methods, or project initiation through project management; there are a number of ways in which potential new software ideas can be brought forwards. The need for software and how this is raised, alongside the four areas outlined above, shall be examined for conversational spaces that may be affected by "Hedging", its uses, and gender.

4.1 The Need for Software

How the need for software or digital intervention is raised will vary between organisations. Some smaller companies may have a more informal model, while other companies may have a systematic process that is followed and applied to all suggestions. More realistically, it could be argued, it is more likely for the truth to lie somewhere between the extremes.

4.1.1 Project proposals

Project proposals may be a way in which ideas are brought forward (Heemstra and Kusters, 2004). A good project proposal should be evidence-based and include due diligence in talking to stakeholders and users, to prove the need for itself. How "Hedging" affects this way of input is seemingly as yet unresearched. Although "Hedging" may or may not be explicit in proposals themselves, the presentations of initial ideas, or presenting the proposals themselves, may be affected by "Hedging", and therefore their likelihood to be accepted may also be affected.

4.1.2 Problem raising

Raising issues with processes, practices or day-to-day working is one of the most common ways in which projects can be raised or changes can occur (Park, 1996; Ashcroft, 2020a). How people are heard when raising problems though, is seemingly one that can quite significantly be affected by "Hedging". For example, if a person says they "think" something can be improved, and another says they "know" it can be improved - it could be argued that it is obvious who will be taken more seriously. This is important to consider as a feminist issue, as we know that men and women use "Hedging" differently (Holmes, 1986), and therefore this is clearly a gendered issue.

4.1.3 Ideation through innovation

Sometimes, one of the main motivators for change can be to utilise existing or upcoming technologies for a new purpose through innovation practices such as through Adobe's Kickbox (Kic, 2021). Innovation, using Adobe's Kickbox (Kic, 2021) involves a six-step process; inception, ideation, improvement, investigation, iteration and infiltration. Each of these stages mainly happens through discussion and therefore, will be affected by language, which may, in turn, be affected by gender. Of course, other frameworks for innovation exist, which opens the opportunity to see how each of these may vary when it comes to the impact that gendered language may have on them.

4.1.4 The effect of gender on the need for software

Regardless of how ideas are brought forward, be it through project proposals, problems being raised, or innovation, many of them are brought forward using some form of language, either digitally or in person. Therefore these are open to being affected by traits of language such as "Hedging", which we know are affected by gender. Therefore, it should be considered, that before the specification process of software creation even begins, that some ideas or changes may already have been lost due to a system that does not allow all members of society to be heard, listened to, and understood. This should be researched further in addition to the above alternative ways of bringing forward the need for software.

4.2 Software Specification

Software specification is the first stage of Software Engineering as defined by Sommerville (2016). How software specifications, or requirements, are gathered varies based on the selected methodology, as well as the organisation or individual in question who is carrying it out. However, most agree that it should include the involvement of all stakeholders, including users of the system. How these are documented can also vary, from requirements listing in Waterfall methodologies (Dima and Maassen, 2018), to User Stories in Agile (Agile Alliance, 2020).

Through both the gathering and recording of requirements, language is instrumental in how this takes place. If the stakeholder engagement is done through *written formats* such as emails, instant messaging or surveys, there are many ways in which language will affect the interaction. If they are done through *spoken interaction*, e.g. video calls, audio calls, or in-person meetings, this also leaves room for misinterpretation, the personal preferences of individuals overriding, or what is recorded as the correct process i.e. human error. People are not without their bias, however unconscious.

However, it could be argued that it is not just the perception of these interactions that will affect the specification stage, but how current issues or ideas are raised. If "Hedging" is used, then perhaps those who use it, often women (Holmes, 1986), may be less likely to be heard and understood. What should be researched further is a deeper understanding of the effect of "Hedging" on this stage, and if any negative causation is found, what can be done to mitigate the effects of this.

4.3 Software Development

Software Development is the second stage of Software Engineering as defined by Sommerville (2016), and is often carried out by men due to the in-balance in the field, with just 11% of software developers being women (Criado Perez, 2019).

Beginning with the interpretation of the requirements, this may be affected by language in both how they have been written, and how they are understood, and as, again, this is done through language, this must also be researched further to find any potential areas of difference regarding gender. It is the subtleties here which could prove telling.

Regarding the programming itself, how men and women code and approach the problem has also previously been shown to be different (Terrell et al., 2017; Ashcroft, 2018), but how this may be affected by "Hedging" remains to be seen. Although language will have an impact on their interactions throughout the process (Ashcroft, 2018), for example in discussing how a problem or requirement should be approached, if and how the code itself differs due to a direct impact of "Hedging" remains to be seen. Previous work in this area does however indicate that approach in the code will differ based on gender and single sex groups (Ashcroft, 2018), the examination of differences when mixed groups are observed could also prove interesting.

Although existing literature shows that development is affected by gender, the specificity of the effect "Hedging" has on this remains to be seen, and definitely requires further investigation.

4.4 Software Validation

Described as the third stage of SE, Software Validation is vital to ensure the software that has been created meets the requirements outlined at the beginning of the process (Sommerville, 2016) (see Section 4.2), therefore, not only must this

step allow for equal representation and equity, but it relies on the same being said for the specification stage. Though testing takes many forms, some of which can be automated, there is also the need to involve as diverse a range of testers, as there are users. Examples of where this has not been the case are varied and ever prevalent in both business and the media, e.g. an Apple Watch battery not lasting for customers with darker skin tones (BBC, 2015).

When it comes to how language can affect the Software Validation stage, how users are asked, and listened to in their responses should be carefully planned, considered and analysed. It is still, of course, important that testers are listened to and comments noted if there is a problem with the tested software, but also that their comments are taken seriously regardless of whether "Hedging" is used. This may of course, not be the case, but again further research is needed to uncover any differences.

4.5 Software Evolution

The stage of Software Evolution recognises the need for continued development and support of systems once they are built and implemented (Sommerville, 2016). This software being correct and appropriate for use, is of course dependent on the correct specification, development and validation; but this stage may also leave room for a lack of input, or a misinterpreted output from all stakeholders. This could, for example, reflect similar practices to problems being raised, and the rest of the process being cycled through; and therefore all the above potential implications of "Hedging" once again apply.

4.6 The Potential Effect of "Hedging"

The overall process of ideation through to implementation involves discussion and the use of language throughout. Although the amount of discussion may vary from stage to stage, and this may also vary between organisations, what is consistent in practice and literature, is that this does occur. With existing research on language, gender, and "Hedging", it seems that there is a gap in the literature when it comes to how these areas affect the Software Engineering process in its entirety. Although literature has begun to emerge on its effect on the design process (Ashcroft, 2021), the rest of the process, from the need for software to its evolution, remains to be researched further with regards to gendered language. The effect of "Hedging", or other existing known gendered differences in language, should be examined alongside each process within Software Engineering to fully uncover "Hedging" as a feminist issue in Computer Science (as discussed further in Section 6).

5 Discussion

Software development, when done well, is a process in which designers, developers, stakeholders and users are involved throughout the entire process.

Therefore it stands to reason equal representation is crucial in the designing and creation of software, as the users of the software themselves are often a diverse range of people. However, it is clear that representation and being in the room is simply not enough. There is very little reason to have representation in the room, if they are not heard, respected, or even given the opportunity to speak. The role "Hedging" plays in this is one of vital importance, and is neither the fault of the individual men or women in the room, but a product of how they were all raised- for example men are significantly more likely to express opinions as certainty, in comparison to women (Holmes, 1986). The societal and systemic raising of children into adults in this way continues to perpetuate a cycle where men feel they are more confident in their opinions, and more likely to express them as facts; and women are much more likely to use "Hedging" as a linguistic tool to express uncertainty, perhaps when there is none, as a means of *taking up less space*.

Therefore, the impact that this has on software, the tools used by many in their day-to-day lives, as part of social interactions, work and for countless other reasons, cannot be understated. If women in design meetings are not listened to, due to the use of "Hedging" or otherwise, how can it be guaranteed that the software, or even hardware, that is created is appropriate for use? There are countless examples of technology not being suitable for wider groups making it into the market, with a release of the aforementioned Apple Watch's battery life not lasting for those with pigmented skin (BBC, 2015), or phones being designed for the average size of men's hands and not women's (Criado Perez, 2019). These are products designed by large companies with the budget and resources to ensure that these mistakes are not made, and yet consistently they are. Which leads this author to assume it is not a lack of ability, but a lack of willingness. With men being more likely to achieve positions of leadership, regardless of past experience (Reuben et al., 2012), how can it be ensured that their best interests are to support those who do not reflect their experience of life? It could, of course, be argued that empathy is a means of overcoming this, and by being able to place themselves in the shoes of others they would be able to see what is needed from the perspective of all potential users, which could also be argued to be the trait of simply a good designer. However, this only works if all information is openly and willingly shared and known, a designer does not know what they don't know e.g. if a designer does not know that on average women's hands are smaller than their own men's hands, how would they then practice empathy in order to change their design?

The above makes the case for representation in the room but does not discuss what happens once equal representation is present, the situation is arguably only slightly improved if when people of varying genders, races and backgrounds are in the room, and will only be significantly improved once they are listened to. As discussed in Section 4, the effect "Hedging" can have on each stage of the Software Engineering process is potentially quite large, but needs to be researched further to be understood on a more complex level. Furthermore, when considering Feminist Methodologies, as discussed in Section 2, these must also be reflected within the scope of Language and HCI more specifically. Whilst this paper has presented a

theoretical understanding of areas for further research, it is vital that this is carried out, as outlined below.

6 Further Research

Though this paper has examined the existing literature surrounding "Hedging", Feminist CA and Software Engineering, there is still a clear need for further research in a number of areas. Therefore, this author recommends five areas for future research.

1. "Hedging" and the need for software. To what extent, if at all, does "Hedging" affect the need for software being raised? As discussed in Section 4.1, could the impact of "Hedging" on discourse effect which problems or ideas for development are brought forwards and how they are perceived? A large scale study of organisations of varying sizes, over time, would be needed to uncover this and best practices if any problems are found.

2. "Hedging" and software specifications. The impact "Hedging" has on software specifications being created, including both the miscommunication in the recording of requirements and their interpretation (as discussed in Sections 4.2 and 4.3) should be examined in closer detail. If specifications are not accurately recorded due to the gender of those raising them, or if they are misinterpreted when the development begins, this is definitely a gendered and feminist issue, and should therefore be investigated more thoroughly.

3. "Hedging", design and coding. In the software development stage (see Section 4.3), there are a number of areas that could be affected by "Hedging". Coding differences, when an analysis of GitHub pull requests was carried out, have been found when it comes to gender (Terrell et al., 2017) and whether this has any link to "Hedging" has yet to be investigated. It could be that these areas are not at all linked, but this should be researched to know for certain and to allow for any subtleties to be examined in closer detail. Furthermore, when it comes to the UX design, which takes place over the specification and development stages, how is this impacted by "Hedging"? Although some research has started to develop in this area (Ashcroft, 2021), it has only been done on a small scale and not in active development or design teams, this should be looked at in practice.

4. "Hedging", validation and evolution. When it comes to software validation/testing (see Section 4.4) it is important to consider; does "Hedging" affect how seriously comments are taken? Even under the assumption that there is a wide range of testers or people carrying out the validation. Then, assuming representation across the board is present, are participants' issues or concerns addressed? Furthermore, on the topic of fair representation of users in the room, there seems to be a landslide agreement in the CSCW and HCI community that this is important, but there is seemingly no framework in place to support this that can be adopted by organisations. In addition to this, as discussed above, this should be done alongside the provision of training and support regarding compassion and empathy, so that validation is not the only means of ensuring

products are fit for use, but one of many. Therefore if one means of overcoming inequality should fail, marginalised groups (including underrepresented genders) are not further disadvantaged. When it comes to software evolution (as discussed in Section 4.5), an area for further research comprises of if "Hedging" increases the misinterpretation when evolution is discussed, and as raised above, in all other stages.

5. "Hedging" and methodologies. Finally, when it comes to all of these above areas of further investigation, how appropriate are the existing methodologies being practiced to uncover the effect of gender on the Software Engineering Life-cycle? An example of where Feminist Methodologies have been applied to an existing methodology is that of Conversation Analysis and the work done by Stokoe (2004). Taking an existing practice, or designing a new one, and ensuring it does not contain any existing bias from preexisting potentially patriarchal research, is vital to make sure that any further research done in the area of Computing, is not perpetuating this (as discussed in Section 2).

7 Contribution and Conclusion

What needs to be considered in more detail within HCI and CSCW is the subtleties of social interaction and the impact they will have on the entire development process. Similar to the work of Stokoe (2004) on Feminist Conversation Analysis, this needs to be applied within the realm of Computing, with enough attention given to the detail of not only *what* is being said, but *how* it is being said when a conversation takes place. This paper has dissected each component of the Software Development Life-cycle from a feminist standpoint and understanding of feminist linguistics, and poses that a more subtle version of this is required; i.e. it is not really what people say, but how they say it.

Theoretically, this paper contributes by taking CA theories, such as interruptions and transitions, and considers these through the lens of existing Feminist Methodologies, such as the consideration of power dynamics. Taking this example further, CA has been known to consider power but never has taken this further because power is not the focus (Button and Sharrock, 2016). The question that should be asked is, how can these power dynamics lead into gender in this way? The practical focus on whether women, and other under-represented genders, have something to say, and the means in which they say it and this is perceived, will have an impact on the products created. The subtlety discussed in this paper will have implications for theory, and therefore further research into the effect gender has on the need for software, specification, design and coding, validation and evolution, and methodologies must all be carried out with these nuances in mind.

Acknowledgments

With thanks to my PhD supervisors Dr Mark Rouncefield and Dr Lynne Blair for their support and guidance, and the reviewers for their comments on structure. I would also like to thank Dr Emily Winter for first bringing the concept of "Hedging" to my attention several years ago.

References

- (2021): 'Kickbox Foundation'.
- (2022): 'CSCW Theories and Concepts'. EUSSET Colloquium.
- Agile Alliance (2020): 'What Is Agile Software Development?'.
- Ashcroft, A. (2018): 'A nice brain teaser'. In: *ACM International Conference Proceeding Series*.
- Ashcroft, A. (2020a): 'Gender Differences in Innovation Design'. *OzCHI '20, December 02–04, 2020, Sydney, NSW, Australia*.
- Ashcroft, A. (2020b): "'Hedging' and Gender in Participatory Design", no. 1986, pp. 176–180.
- Ashcroft, A. (2021): 'Do I belong here? An exploration of meeting structure and language, alongside gender and a sense of belonging.'. *OzCHI '21, November 30-December 2, 2021, Melbourne, VIC, Australia*.
- Barbour, K. (2018): *Embodied Ways of Knowing: Revisiting Feminist Epistemology*, pp. 209–226. London: Palgrave Macmillan UK.
- BBC (2015): 'Tattooed Wrists Can Stop Wearables Like The Apple Watch Working'.
- Bellini, R., A. Strohmayer, E. Alabdulqader, A. A. Ahmed, K. Spiel, S. Bardzell, and M. Balaam (2018): 'Feminist HCI: taking stock, moving forward, and engaging community'. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. pp. 1–4.
- Benwell, B. (2006): *Discourse and identity*. Edinburgh, [England]: Edinburgh University Press.
- Blackwell, E. C. (2020): 'Taking Up Space: A Phenomenological Study of the Shared Experiences of Black Women in STEM Graduate Programs at Predominantly White Institutions'.
- Boden, D. (1994): *The Business of Talk: Organizations in Action*. Cambridge: Polity Press.
- Button, G. and W. Sharrock (2016): 'In support of conversation analysis' radical agenda'. *Discourse Studies*, vol. 18, no. 5, pp. 610–620.
- Catolino, G., F. Palomba, D. A. Tamburri, A. Serebrenik, and F. Ferrucci (2019): 'Gender diversity and community smells: insights from the trenches'. *IEEE Software*, vol. 37, no. 1, pp. 10–16.
- Criado Perez, C. (2019): *Invisible Women: Exposing Data Bias in a World Designed for Men*. Vintage, 10 edition.
- Davis, C. (2019): 'Feminist Rhetorical Practices in Digital Spaces'. *Computers and Composition*, pp. 132–141.

- De Hertogh, L. B., L. Lane, and J. Ouellette (2019): “Feminist Leanings:” Tracing Technofeminist and Intersectional Practices and Values in Three Decades of Computers and Composition’. *Computers and Composition*, no. 51, pp. 4–13.
- Dima, A. and M. Maassen (2018): ‘From Waterfall to Agile software: Development models in the IT sector, 2006 to 2018. Impacts on company management.’. *Journal of international studies*, pp.315–325.
- Dixon, J. and D. Foster (1997): ‘No Title’. *Journal of Psycholinguistic Research*,, vol. 26, no. 1, pp. 89–107.
- Flax, J. (1993): ‘Disputed Subjects: Essays on Psychoanalysis, Politics, and Philosophy’.
- Ford, C. (2008): *Women Speaking Up: Getting and Using Turns in Workplace Meetings*. New York: Palgrave.
- Friedrich, M. C. G. and E. Heise (2019): ‘Does the Use of Gender-Fair Language Influence the Comprehensibility of Texts?’. *Swiss Journal of Psychology*.
- Hancox-Li, L. and I. E. Kumar (2021): ‘Epistemic Values in Feature Importance Methods: Lessons from Feminist Epistemology’. In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. New York, NY, USA, p. 817–826, Association for Computing Machinery.
- Heemstra, F. J. and R. J. Kusters (2004): ‘Defining ICT proposals’. *Journal of Enterprise Information Management*.
- Holmes, J. (1986): ‘Functions of You Know in Women’s and Men’s Speech Language in Society’. *Source: Language in Society*.
- Holmes, J. (1990): ‘Hedges and boosters in women’s and men’s speech’. *Language and Communication*.
- Hyland, K. (1996): ‘Writing without conviction? Hedging in science research articles’. *Applied Linguistics*, vol. 17, no. 4, pp. 433–454.
- Jay, A. (2009): *How to Run a Meeting*, Harvard Business Review Classics. Harvard Business Review Press.
- Kitzinger, C. (2008): ‘Developing feminist conversation analysis: A response to Wowk’.
- Leau, Y. B., W. K. Loo, W. Y. Tham, and S. F. Tan (2012): ‘Software development life cycle AGILE vs traditional approaches’. In: *International Conference on Information and Network Technology*, Vol. 37. pp. 162–167.
- Liddicoat, A. J. (2011): *An Introduction to Conversation Analysis: Second Edition*. Continuum International Publishing Group.
- Murphy, B. (2010): *Corpus and sociolinguistics : investigating age and gender in female talk*. Amsterdam ; Philadelphia: John Benjamins Pub. Co.
- O’Leary, Z. (2017): *The essential guide to doing your research project*. SAGE, 3rd edition.
- Park, D. (1996): ‘Gender role, decision style and leadership style’. *Women in Management Review*, vol. 11, no. 8, pp. 13–17.

- Paulus, T., A. Warren, and J. N. Lester (2016): 'Applying conversation analysis methods to online talk: A literature review'. *Discourse, Context Media*, vol. 12, pp. 1–10.
- Reuben, E., P. Rey-Biel, P. Sapienza, and L. Zingales (2012): 'The emergence of male leadership in competitive environments'. *Journal of Economic Behavior and Organization*.
- Sacks, H., E. A. Schegloff, and G. Jefferson (1974): 'A Simplest Systematics for the Organization of Turn-Taking for Conversation'. *Language*, vol. 24, no. 2, pp. 111–134.
- Schlesinger, A., W. K. Edwards, and R. E. Grinter (2017): 'Intersectional HCI'.
- Sommerville, I. (2016): *Software engineering*. Boston, [Massachusetts]: Pearson, tenth edition, global edition. edition.
- Speer, S. A. and E. Stokoe (2011): *Conversation and gender*. Cambridge: Cambridge University Press.
- Stokoe, E. (2018): *Talk: The Science of Conversation*. Robinson.
- Stokoe, E. H. (2004): 'Gender and discourse, gender and categorization: Current developments in language and gender research'. *Qualitative Research in Psychology*, vol. 1, no. 2, pp. 107–129.
- Stokoe, E. H. and J. Smithson (2001): 'Making gender relevant: Conversation analysis and gender categories in interaction'. *Discourse and Society*, vol. 12, no. 2, pp. 217–244.
- Svennevig, J. (2012): 'Interaction in workplace meetings'. *Discourse studies*, vol. 14, no. 1, pp. 3–10.
- Tannen, D. (1994): *Talking from 9 to 5*. New York: Avon Books.
- Terrell, J., A. Kofink, J. Middleton, C. Rainear, E. Murphy-Hill, C. Parnin, and J. Stallings (2017): 'Gender differences and bias in open source: Pull request acceptance of women versus men'. *PeerJ Computer Science*.

Riccardo Bovo, Daniele Giunchi, Enrico Costanza, Anthony Steed, Thomas Heinis (2022): Shall I describe it or shall I move closer? Verbal references and locomotion in VR collaborative search tasks.. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_ep02

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_ep02

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

Shall I describe it or shall I move closer? Verbal references and locomotion in VR collaborative search tasks.

Riccardo Bovo, Daniele Giunchi, Enrico Costanza, Anthony Steed,
Thomas Heinis

Imperial College London, University College London
rb1619@ic.ac.uk, d.giunchi@ucl.ac.uk

Abstract. Research in pointing-based communication within immersive collaborative virtual environments (ICVE) remains a compelling area of study. Previous studies explored techniques to improve accuracy and reduce errors when hand-pointing from a distance. In this study, we explore how users adapt their behaviour to cope with the lack of accuracy during pointing. In an ICVE where users can move (i.e., locomotion) when faced with a lack of laser pointers, pointing inaccuracy can be avoided by getting closer to the object of interest. Alternatively, collaborators can enrich the utterances with details to compensate for the lack of pointing precision. Inspired by previous CSCW remote desktop collaboration, we measure visual coordination, the implicitness of deixis' utterances and the amount of locomotion. We design an experiment that compares the effects of the presence/absence of laser pointers across hard/easy-to-describe referents. Results show that when users face pointing inaccuracy, they prefer to move closer to the referent rather than enrich the verbal reference.

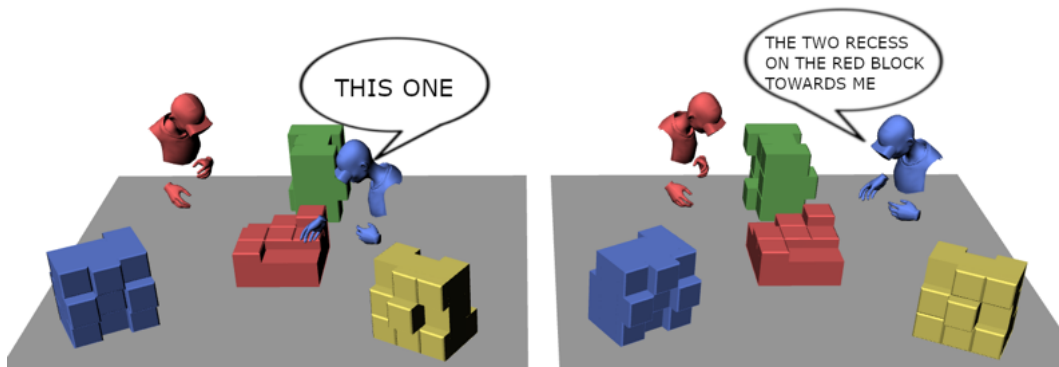


Figure 1: On the left an example of implicit verbal reference aided by a hand pointing action at a close distance from the referent. On the right, the equivalent reference is aided by a more detailed verbal description of the referent but lacks the hand pointing action from a close distance.

1 Introduction

Immersive collaborative virtual environments (ICVE) with user embodiment (i.e., avatars) support collaboration by providing a shared setting where collaborators have a sense of each other's presence (Benford et al., 1995). In ICVEs, the user's embodied hands behaviour is a non-verbal cue that integrates verbal communication during collaboration (Hindmarsh et al., 1998). For example, users can point to a referent during an utterance to trigger mutual orientation and visual coordination Moore et al. (2007). Hand pointing in conjunction with verbal, spatial references is called deictic pointing. Previous studies explore deictic pointing with distant targets (from fixed distances), measuring the accuracy of different hand pointing supports (Mayer et al., 2018, 2020; Wong and Gutwin, 2014, 2010). Outcomes from previous studies highlight how the degree of precision needed for the pointing gestures depends on how complex it is to describe the referent using utterances (Wong and Gutwin, 2014, 2010). However, in modern ICVE, users might get as close as needed to the referent and adapt to the accuracy required to perform the pointing gesture. Therefore, when faced with a lack of accuracy, will users spend time adjusting their distance from the target or overcome the difficulties of describing the target?

While in a physical environment, it is not always possible to move closer to an object of interest, in an immersive ICVE, this is not a problem as there are no physical barriers. In such scenarios, users can avoid inaccurate distance pointing by moving closer to the referent. However, the movement has a temporal cost: the time required to move closer to the point of interest. As Wong and Gutwin (2014) highlight, another approach consists of users enriching their verbal references with enough details to compensate for the pointing gesture's lack of precision. On the other hand, such a verbal supplement comes with a higher temporal and cognitive cost for both the performer and the reference recipient Wong and Gutwin (2014); D'Angelo and Begel (2017). Previous studies define pointing accuracy as a

function of distance from the referent (Mayer et al., 2020; Wong and Gutwin, 2010). Accurate pointing can be performed from a far distance with a laser pointer for support or performed without a close distance to the referent. Inaccurate pointing consists of users who do not use/have laser pointers from a distance and choose to compensate with explicit verbal references.

The research community established the importance of laser pointers to achieve accurate pointing, but like any other tool or metaphor of interaction, laser pointers can be included or not in an ICVE. Some reasons for not including pointers can be the following: data visualisation issues such as a hidden or occluding cursor (especially if the informative area is dense), many users with cursors, and noise-induced by body jittering in high-density information areas Batmaz and Stuerzlinger (2019).

This study explores the trade-off between using locomotion to approach the referent or the alternative use of explicit verbal references to deal with the lack of pointing accuracy. We look at how this trade-off varies across conditions of lack/availability of laser pointers and conditions related to how complex/easy it is to describe the various referent in the scene (Figure 1). We explore such trade-offs in the context of visual search tasks, which are recognised as a proxy for many other tasks performed synchronously by pairs of participants in ICVE's Prilla (2019).

We run an experiment with 20 participants quantifying implicit/explicit references, locomotion and, in addition, visual coordination, which is highly correlated to the quality of pointing-based communication (Schneider and Pea, 2013). We use two datasets of different complexity representing two levels of difficulty in describing the referent: a simple puzzle and a very detailed 3D satellite map. In the simple 3D puzzle, pieces can be described by colours or labels, while on the map, places need to be referenced via 2D coordinates, which requires a greater cognitive effort. Inspired by previous CSCW work D'Angelo and Begel (2017) we measure the number of implicit/explicit deixis and the number of successful/unsuccessful deixis. Moreover, we measure users' movement in the space and task performance (task score and completion time).

The data collected shows statistically significant differences in locomotion performed when distance-pointing support is unavailable. Both data and observations confirm that when users lack support for distance pointing, they prefer to move closer to the referent to perform accurate pointing gestures rather than formulate a more complex verbal reference. We can see this change no matter the complexity of the task. The data collected also shows a statistically significant increase in visual coordination when laser pointers are available, which confirms previous work Moore et al. (2007).

Our results enable designers to understand how different elements (embodiment, locomotion, laser pointers) available in immersive ICVE impact pointing-based communication during a generic collaborative visual search task. Thus, our work can contribute to a more proficient interaction by outlining design implications. Presence of locomotion and the freedom for the user to move throughout the whole environment remove the need for distance pointing support.

In this way, an efficient locomotion system increases the rates of proximal pointing instead of promoting a cursor for distal pointing. However, laser pointers support may need to be considered if the collaborative task requires high visual coordination. Our study thus helps to make informed choices when designing an ICVE.

2 Related work

Pointing-based communication is ubiquitous in collaborative work. Within physically co-located scenarios, a pair of collaborators may use their hands and voice to engage in pointing-based communication. For example, indicating an object of interest by pointing hands towards it during an utterance is a common interaction called deictic pointing or deixis. During deixis, the interlocutor (i.e., recipient of the deixis) has to mentally project the collaborator's hand directly onto the observed scene to understand the referent of the deixis (i.e., understand the target object) (Higuch et al., 2016; Pfeiffer et al., 2008; Wong and Gutwin, 2014).

Pointing-based communication, however, can also be supported by laser pointers. A pointer's spotlight projected onto the observed scene allows identifying the referent unambiguously (Hindmarsh et al., 1998). Additionally, it facilitates the interpretation of the pointing gesture by removing the cognitive effort of projecting the hand/head directly onto the observed scene. Using a laser pointer might avoid any incorrect mental projection or ambiguous unclear projection results. Essentially pointers increase the awareness, during deixis, of a collaborator's visual focus (Piumsomboon et al., 2017).

Pointing-based communication is possible in co-located scenarios and remote scenarios thanks to either embodiment (i.e., avatars) and enhanced behaviour (i.e., pointers). There are several examples of remote collaboration scenarios in which pointing-based communication is possible, to mention a few: remote pair programming (D'Angelo and Begel, 2017), support of local workers by remote experts (Bai et al., 2020), remote collaboration in immersive VR environments (Moore et al., 2007).

2.1 Pointing-based communication in remote desktop collaborations

Pointing based communication can occur as long as collaborators have the means to point towards an object of interest while also communicating verbally. For example, several studies investigate pointing-based communication using gaze pointers (i.e., enhanced behaviour of eyes) in the context of 2D desktop remote collaboration (Villamor and Rodrigo, 2018; Jermann et al., 2011; Nüssli, 2011; Pietinen et al., 2008).

These studies show how visual aids based on the eye-tracked behaviour of collaborators (i.e., gaze-pointers) increase mutual awareness of visual focus, higher visual coordination and better collaboration quality. Schneider and Pea (2013) explore how depicting gaze in a remote desktop collaboration of two users

performing a visual task increases visual coordination and enhances visual collaboration quality. When visual aids, such as pointers, are used, collaborators look at the same objects at the same time more often than without visual aids. Additionally, such increased visual coordination seems to aid communication about the visual context. For example D'Angelo and Begel (2017) explore visual aids (based on real-time eye-tracked behaviour) and prove that such visual aids improve communication by reducing the number of explicit utterances during deixis.

However, findings from the 2D desktop environment remote programming and visual analysis do not generalize to the immersive VR environments because the reviewed scenarios lack embodiment and locomotion (both elements present in state-of-the-art immersive VR collaboration environments). Embodiment, especially hand representation and hand real-time tracking behaviour, is the natural behaviour used in deictic pointing. However, in 2D desktop environments, the gaze is used as an input for pointing. While gaze can be thought of as coinciding with visual attention, it is a behaviour that is less deliberate and thus less controllable than the behaviour of hands. A second significant difference is related to fragmentation (Wong and Gutwin, 2014; Hindmarsh et al., 1998), or in other words, the fact that large parts of the environment in VR are not visible to the users, unlike the 2D desktop screen is. Fragmentation impacts pointing-based communication because the pair of collaborators may not be seeing the same subset of the 3D environment during deixis. They may thus not be able to see the collaborators' embodiment or the pointing visual aid. Moore et al. (2007) highlights how the observability of embodied activity and the projectability of gestures are essential aspects of pointing-based communication. While 2D desktop remote programming work may inspire metrics such as visual coordination and implicitness/explicitness of deixis utterances, their results are not necessarily generalizable to immersive VR collaboration.

2.2 Pointing-based communication in ICVE

Finally, ICVE offers the same degree of embodiment of mixed reality scenarios. Real-time tracked behaviour of hands/head allows natural pointing behaviour and natural exploration of the scene via head movements and locomotion. Several immersive VR studies explore the accuracy of hand pointing gestures. Mayer et al. (2018) propose adaptations to hand pointing in immersive VR that enhance the precision and accuracy of the pointers representations through spatial distortion. Mayer et al. (2020), in a similar way to Sousa et al. (2019) explores the approaches to improve precision by warping gestures to adjust pointing to the target.

However, while these recent studies aim to improve hand pointing accuracy, they do not evaluate the effect that pointers have on collaboration focusing only on the quantification of the pointing accuracy. All these works measure the accuracy of pointing from fixed distances, avoiding any form of locomotion within the scene. Our work aims to fill this gap, introducing specific tasks where we require

the participants to move freely in the scene. An additional study from Bai et al. (2020) proposes a remote collaboration system that introduces an asymmetric interaction between a VR user and an AR user sharing a live 3D panorama of their surroundings. Differently from this study, our VR system provides both symmetric interaction and interface, and we focus on measuring the impact of locomotion on pointing-based communication.

2.3 How users compensate for inaccuracies during distance pointing

Previous studies explore techniques to improve accuracy and reduce errors when hand-pointing during pointing-based communication in immersive collaborative virtual environments (CVE). However, in a CVE in which users can move (i.e., locomotion), distance pointing (and its negative consequences) can be easily avoided by users' choice of increasing proximity to the referent. Additionally, a user could choose to compensate for imprecise distance pointing by enriching (adding details) to a verbal reference during a pointing gesture.

In an immersive CVE with embodiment and locomotion, we compare the presence and absence of pointers to understand if and how users compensate to avoid pointing errors and lack of precision. We also use several quantitative measures to understand how behaviour changes impact the quality of pointing-based communication. Inspired by previous CSCW remote desktop collaboration, we identify three easily quantifiable metrics: visual coordination, the implicitness of deixis' utterances, and references' success. Such metrics represent the quality of pointing-based communication during a collaborative task.

Previous literature allows us to define accurate pointing (both from the points of view of the producer and observer) and inaccurate pointing. Pointing gestures can be either proximal or distal Schmidt (1999). When indicating proximal referents, the producer of a pointing gesture can touch the target, and observers can identify targets with confidence Bangerter and Oppenheimer (2006). Therefore, consider proximal pointing is considered accurate as there is no room for misinterpretation.

With distal pointing, the observer needs instead to extrapolate the vector direction defined by the pointer's posture Bangerter and Oppenheimer (2006); Batmaz and Stuerzlinger (2019). However, previous studies have found that using a cursor improves mid-air pointing precision thanks to visual feedback and removes the need to extrapolate the direction of the pointing gesture again thanks to the visual depiction of the cursor Mayer et al. (2018). Therefore, we consider distal pointing with the cursor accurate as there is no room for misinterpretation, while we define distal pointing without the cursor as inaccurate.

While previous works offer several methods to improve the accuracy of pointing via machine learning models in our study, we explore how users deal with the lack of accuracy in an ecological context, in particular, related to visual analysis tasks.

3 Study Design

In the following subsections, we detail different aspects of the experiment. This study has been approved by the UCL Interaction Centre (UCLIC) Research Department's Ethics Chair.

3.1 Participants

Twenty-four participants (twelve pairs) volunteered to take part in the remote study. The data of two pairs of participants was used to pilot the study and test the application, while the remaining ten pairs were used for the data analysis. One condition during recruitment was for participants to own or have access to the specific VR HMD: Oculus Quest. This condition was because the experimental session was conducted remotely via teleconference software and then via the VR application. Participants were recruited online via forums and social networks groups dedicated to the Oculus Quest headset and Slack channels dedicated to HCI VR research participant pooling. Participants were recruited individually and then matched up in pairs based on their time availability to conduct the experiment. All participants provided informed consent and received £15 compensation for participating. For the study, pairs of participants were asked to work together in a remote collaborative visual analysis task. Participants were familiar with VR devices as they owned or had access to a HMD's headset. All participants had at least a university grade (6 PhD Candidates, 8 PhDs, 6 MScs, 4 BAs). The mean age was 33 years old with a standard deviation of 8.3. The 88% of the participants was male, and the 12% female.

3.2 Setup

To keep the application development simple and to avoid noise due to differences across VR HMDs, we decided to target a single device for the experiment. The selected headset (Oculus Quest) is 6 degrees of freedom (DoF) untethered VR HMD, with a 60 Hz refresh rate. We chose this headset because of its popularity and low retail price. We developed an application for collaborative visual analysis of 3D data using Unity (version 2018.4.14f1) and the Oculus unity SDK. The application enables the visualisation of different types of 3D data sets (i.e., terrains, 3D networks, CAD files). The application enables each participant to join a real-time session in which other participants' presence is represented by avatars (i.e., Oculus Avatar SDK) as shown in Figure 2. Each participant in the VR space is free to move in any direction using a thumbstick controller or physically move using the 6 DoF of the VR HMD. Avatar movements are streamed via the network, so their behaviour (head and hand movements) and position in the virtual space is reproduced with low latency. The application also enables participants to talk to each other using the embedded microphone and speakers of the VR HMD. Additionally, the setup supports an observer/moderator to be present in the VR session and environment.

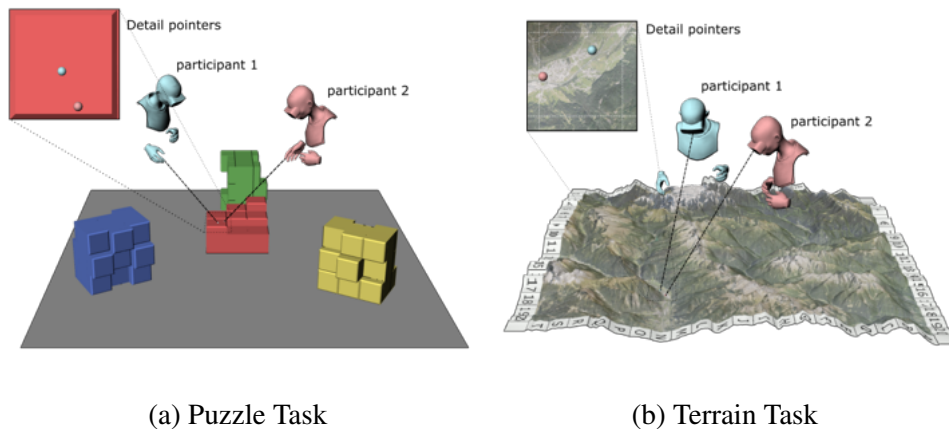


Figure 2: A pair of participants collaborate on the visual analysis tasks in the 3D environment. a) Participants are using a hand pointer while performing a four-part 3D puzzle. b) Participants identify the four largest settlements in a terrain dataset using a hand pointer. The hand direction visualised as a series of dotted lines is displayed in the image only to illustrate the difference between head and hand pointers. Both task environments have a size of 3x3 meters.

3.3 Pointers

The pointer consists of a small (1cm) sphere depicted at the intersection between the direction of the hand and the visualised data. Hands are tracked via controllers, and the hand pointer is associated with the dominant hand via the Oculus Unity SDK. The VR HMD tracks the head direction and position. The hand direction, or in other words, the ray departing from the hand, is not visualised; instead, the little sphere is visualised, depicting a small spotlight and therefore displaying the same effect of a laser pointer. The pointers can be seen in the "Detail pointers" window in Figure 2. When the pointer is not present, participants can still point using the hand embodiment as if they were in a physical co-located collaborative scene. The controller triggers approximate the posture of hands, so if a trigger is pressed/released, the correspondent finger is depicted fully contracted or in a straight position. Users can, therefore, intuitively use the index finger to point to referents (Figure 2).

3.4 Experiment Design

We design a 2(pointer)x2(reference difficulty) factors (Table 3a), within-subjects experiment. Participants collaborate on two visual search tasks consisting of identifying visual features in two data sets. The reference difficulty factor consists of two levels: a 3D terrain with hard-to-describe features and a 3D puzzle with an easy-to-describe feature. On the hard level, verbal references can be done using map coordinates or describing features in detail. On the easy level, verbal

references can refer to the colour of puzzle blocks or a unique label number. We argue that the complexity of the features in the satellite map is higher than the simple puzzle geometric shapes to describe and disambiguate. Moreover, map coordinates are more complex to reference than a single puzzle label, as they require users to compose the coordinate by reading both longitudinal and latitudinal labels. Therefore, we argue that the cognitive effort required to describe the map's referent is higher than the puzzle. We validated such a hypothesis by pilots of the experiment. Moreover, experiment results of the number of implicit references further validate this level classification. The pointer factor consisted of two conditions: a condition without any pointer and a hand laser pointer, as previous work validates pointers as successfully supporting pointing based communication Moore et al. (2007).

3.5 Task

The two tasks are collaborative visual search tasks. Visual search task is considered a proxy for many other tasks to be done together in VR synchronously, which include finding virtual objects or information together, jointly referencing the same referent Schmalstieg and Höllerer (2016); Prilla (2019).

For the hard task, we used a scenario common in HCI studies that consist in identifying features on 3D terrain maps. We took inspiration from previous works Šašinka et al. (2019); Liu et al. (2017). 3D terrain data is rich in details. Therefore it is complicated to describe it verbally. In the 3D terrain visual analysis task (i.e. hard verbal reference task), participants must identify the four largest settlements (i.e. cities) and the four largest lakes. The terrain consists of satellite images and elevation extracted from Mapbox, and the coordinates corners in the first dataset are for the top-left latitude 46.56, longitude 11.53 and bottom-right latitude 46.17, longitude 11.92; in the second dataset, the coordinates are top-left latitude 46.62, longitude 10.53 and bottom-right latitude 46.23, longitude 11.92.

For the easy task, we selected a scenario that is very common in collaborative VR tasks: puzzle. For example, many studies can be found in the literature using puzzle quiz Slater et al. (2000); Steptoe et al. (2009); Schroeder et al. (2001); Widstrom et al. (2000); Kim et al. (2014). Such tasks contain a visual analysis component which requires participants to identify compatible blocks by comparing them. In our specific case, we avoided any manipulation to focus on visual analysis and related pointing-base communication. In the 3D puzzle task, users must identify the four puzzle blocks that fit together (2 puzzles were present for each experiment condition). Each block measures 50x50x25 cm, and each of the two sides of the block contains 3x3 puzzle joints. Both puzzle conditions are available to be downloaded from ANON-REPOSITORY. At the beginning of each trial, participants were asked to collaboratively identify and report the four correct features to the experiment moderator. If there was a leading effect (i.e., one participant being the only one active), the experiment moderator would remind the pairs to discuss and agree upon features before reporting them. Both task search

spaces are equal in size and correspond to 3x3 m. The time given to participants is displayed as a countdown on the VR scene and consists of 5 min max for each scene.

3.6 Procedure

At the beginning of each experimental block, participants are given a chance to practise the task and familiarise themselves with sample datasets. The practice time consists of a maximum of 5 min, but participants can interrupt it earlier if needed. The sample dataset used in practice was not used for the task. Users were allowed to train on both the easy (blocks) and hard (map) tasks. During familiarisation, participants can ask questions; this phase ends once both participants confirm understanding the task. Following the familiarisation, participants are asked to perform the task across the two conditions: hand pointer and no pointer. For each of the two conditions, an equivalent variation of each data set is used (two terrains and two puzzles) for four data sets (Table 3b). Trial order and experimental block order were randomised to counterbalance learning effects.

Once participants agree on a feature, they are asked to communicate it to the observer verbally. The observer only acknowledges the communicated data features as recorded if both participants explicitly agree on it; otherwise, the observer prompts a reminder that both participants have to agree. Such constraint forces pairs to work collaboratively. To incentivise engagement with the task, participants are told that if they score above a specific threshold value, they will receive a £15 voucher instead of a £10 voucher (in the end, every participant receives £15 regardless of their score). We recorded audio and video in VR and log position for all the experiment sessions.

		Factor1: Pointer	
		Level 1 No Pointer	Level 2 Hand Pointer
Factor2: Difficulty	Level1 Terrain	No Pointer Terrain	Hand Pointer Terrain
	Level2 Puzzle	No Pointer Puzzle	Hand Pointer Puzzle

Experimental Session Participants Diad					
Experimental Block1 Factor 2 Level 2: Terrain			Experimental Block2 Factor 2 Level 2: Puzzle		
eg	Trial1 F1 L1 No Pointer	Trial2 F1 L2 Hand Pointer	eg	Trial1 F1 L1 No Pointer	Trial2 F1 L2 Hand Pointer

(a) Experiment Design
(b) Experiment procedure

Figure 3: (a) Experiment Design: the experiment has two factors: dataset and pointer. The dataset factor has two levels: 3D surface (terrain), 3D volumes (puzzle). The pointer factor has two levels: No Pointer, Hand Pointer. (b) The experimental procedure is divided in experimental blocks one for each level of the independent variable difficulty, and experimental trials one for each level of the independent variable pointer, plus one trial for task familiarization at the start of each experimental block. Trial order and experimental blocks order were randomised to counterbalance learning effects.

4 Measures

This section gives an overview of the measures collected during the experiment and how we post-process them. We record the head behaviour of both participants. Head gaze is the intersection between the ray starting from the Head position with the direction of Head rotation and the visualised data, which is used to calculate head concurrent pointing behaviour (i.e., visual coordination, section 4.1). Additionally, we record a video/audio stream of the virtual environment for every experimental session of the participants' avatars, containing verbal communication between participants. We use this data to perform implicit/explicit reference analysis (Section 4.3). To understand if the experimental conditions impact temporal and accuracy performances, we also record the task time and task score. Task time is capped to 300 seconds, 5 min to keep the duration of the whole experiment to 20 min max. The maximum number of correct answers for each task is four.

4.1 Visual Coordination

Visual coordination consists of participants' visual focus coupling, or in other words, how well synchronised their visual attention is. As previous work suggests, when users point to a referent during an utterance, this triggers mutual orientation, an essential part of visual coordination. Pointing-based communication is, in this sense, an effort aimed at negotiating shared visual attention during collaborative work Moore et al. (2007). Previous work also explores how visual coordination is highly correlated to the quality of collaboration Schneider and Pea (2013). Therefore visual coordination is a crucial dimension of collaboration in visual search tasks.

The ideal measure of visual coordination would require to use eye-gaze behaviour. However, our study did not use eye-trackers as most low-cost VR HMDs do not have them and running a remote user study during pandemics requires us to target popular low-cost headsets such as oculus quest. Instead, we use head-gaze behaviour, which several studies have reported as a good proxy of eye movements (Biguer et al., 1982; Pelz et al., 2001; Wang et al., 2019). Concurrent head pointing measures the time two participants concurrently point their heads towards the same target simultaneously. For example, when collaborators discuss a visual feature, they are likely to point their heads towards such a feature concurrently. This effect is also described as mutual orientation, identified by Moore et al. (2007) as the first stage of deixis in pointing-based communication.

We post-process head gaze recorded data to measure the time head-gaze overlap during one experimental trial. We define a distance of 20 cm as the threshold for the euclidean distance calculation. Below such threshold, the two head-gaze are considered to point at the same location and above. They are considered to be pointing at different data features. The distance between the two head gaze points

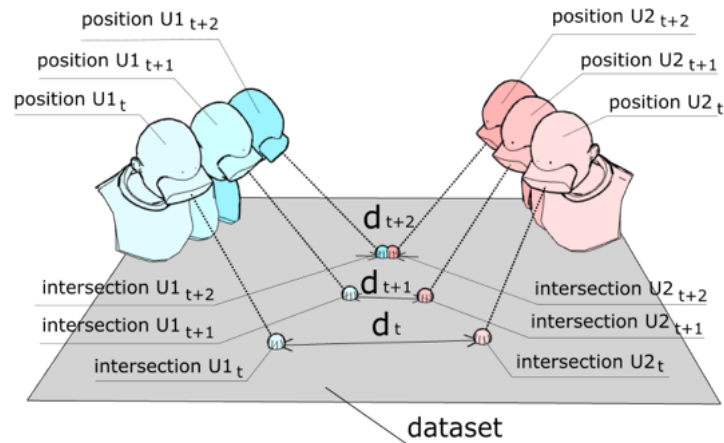


Figure 4: A view of the collected measure of head position, head direction and head signal intersection at a specific moment in time. We post-process head gaze recorded data (intersection) to measure the time head-gaze overlap during one experimental trial by computing the euclidean distance for each time frame. We post-process head position data to measure the time (seconds) participants stay still/move and compare it throughout the different experimental conditions.

is calculated for every sample at time t ; then, we multiply the number of samples by the sampling frequency to obtain the cumulative time of concurrent head pointing (Figure 4).

4.2 Locomotion

When performing an implicit spatial reference (i.e. pointing/utterance) during pointing-base communication, the referent can be misunderstood by the collaborator (i.e., recipient). Such misunderstanding happens because the gesture performer might point imprecisely. Alternatively, the recipient may fail to correctly project the direction of the hand/arm onto the observed scene. A way to improve the accuracy of a pointing action during pointing-based communication consists in moving closer to the referent, so to make sure that the observer/listener won't miss-interpret the direction of the pointing action Wong and Gutwin (2010). Laser Pointers instead allow participants to perform precise pointing. Using a laser pointer, the performer of the pointing action can adjust the cursor position until the cursor lays on the referent, removing ambiguities. Pointers, therefore, allow to perform accurate pointing gestures from a distance (i.e., without having to travel towards the referent) Wong and Gutwin (2014). However, during collaborative visual tasks, participants might be interested in reducing the distance to a referent for other reasons, such as observing it in greater detail or simply increasing its presence by joining a collaborator's working area.

To investigate the impact of locomotion on pointing based communication, we measure how much time each participant spends moving in ICVE during each trial. As part of the experiment guideline, we expressly asked participants to explore the space only via a thumb-stick controller rather than moving physically for safety reasons. Therefore we used the locomotion speed set in the unity environment of 1.6 m/s to determine the ideal threshold to classify intended movement and noise.

We post-process head position data to measure the cumulative time of locomotion and compare it throughout the different experimental conditions. To calculate the locomotion time, we considered only the samples where the velocity is above the threshold of 0.8 m/s, calculating the distance using sampling frequency and velocity and removing small movements and noise.

4.3 Implicit references

Deixis consists of verbal references supported by a pointing gesture. Within a visual search task, deixes are common occurrences as they allow negotiating the collaborative shared visual context.

Deixes can be implicit or explicit: the first requires less information uttered and are also cognitively less demanding D'Angelo and Begel (2017); Wong and Gutwin (2014). Implicit deixis tends to rely more on the accuracy of the pointing action as the utterance does not carry sufficient information to disambiguate the referent. We consider an implicit spatial reference occurring whenever a participant referred to a data feature without explicitly naming any unique property of the object (i.e., name, location, colour). Instead, explicit deixis contains information to disambiguate the referent from the rest of the data set. Such explicit information can consist of: position relative to the user (e.g., on my left/right etc.), object characteristics (e.g., the red block etc.), labels (i.e., a unique textual description) or its absolute position expressed in coordinates (i.e., the data feature in B5).

Understanding how pointing based communication changes when hard-to-describe referents are present, or a lack of distance pointing support means classifying each deixis as implicit/explicit. Such a classification gives us an understanding of how smooth/fast verbal communication is. Additionally allows us to understand the balance with behavioural alternatives, such as getting close to the referent to pinpoint it more accurately.

Inspired by previous CSCW work proposed by D'Angelo and Begel (2017), we transcribed audio of the collected videos and carried out a double-blinded video/text classification of the spatial, verbal references. Two analysts performed the analysis to countereffect the subjectivity of the classification process. If the two interpreters were unclear if an instance was implicit or explicit, they conducted a collaborative post-analysis to reach convergence.

We also classify each reference as successful/unsuccessful. Such classification allows us to understand if and how locomotion impacts the effectiveness of point-based communication when there is a lack of support for distant pointing. A

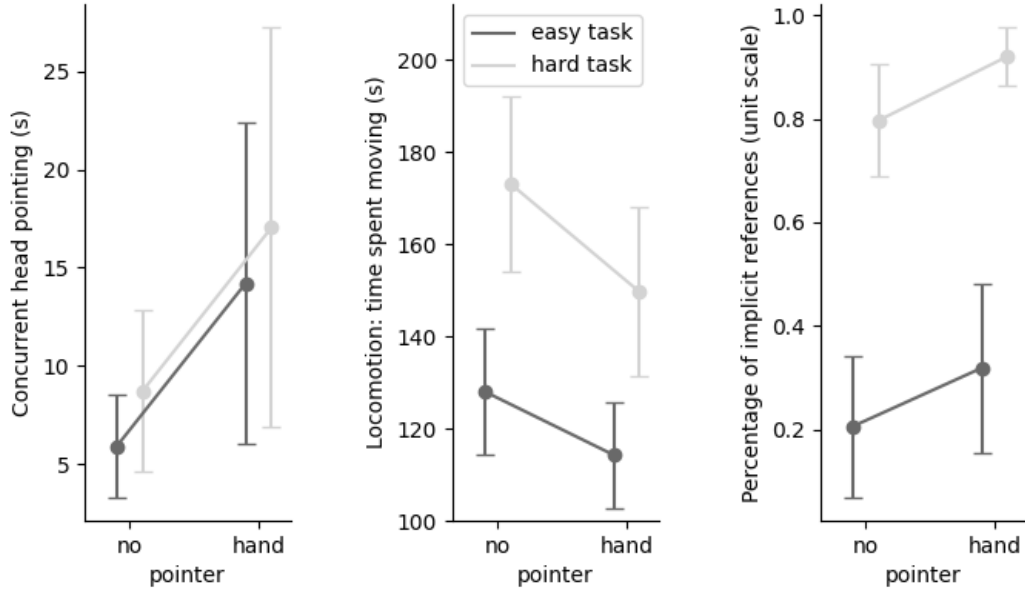
reference is considered unsuccessful when the recipient misinterprets the correct referent or if the recipient ignores the deixis.

5 Statistical Analysis

We performed a repeated measure ANOVA test (using JASP) on the data we collected and post-processed. For the measures of temporal and accuracy performances and the number of unsuccessful deixes, the analysis did not return any significant difference across conditions. For these measures, we won't report the analysis result for conciseness. Our results of visual coordination are achieved by a set of 10 samples (10 pairs of participants). While for locomotion and implicit references, all 20 participants are measured individually, thus equivalent to 20 samples.

5.1 Visual Coordination

The 2 way ANOVA analysis results show one main effect related to the factor: *Pointer* p-value $<.001$ (Table I and Figure 5a). When participants have a laser pointer, they spend approximately 8 seconds more pointing their head towards the same data subset. To contextualize this measure, the average duration of a task is 230 sec, representing approximately 3.4% of the time. However, from observations, we can see that the task time is split between independent work (scanning data visualization independently) and collaborative work (discussing the interpretation of data features). Considering that visual coordination only relates to collaborative work, we argue that the 3.4% of time represents a much higher value within the collaborative stages.



(a) Visual Coordination

(b) Locomotion

(c) Implicit references

Figure 5: Descriptive plots: on the horizontal axes the pointer conditions, on the separate lines the difficulty of explicit references (i.e., hard task and easy task), error bars display the confidence interval of 95%.

Table I: ANNOVA: Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p
(a) Visual Coordination					
pointer	1001.618	1	1001.618	22.919	< .001*
difficulty	11.694	1	11.694	0.114	0.743
pointer * difficulty	29.941	1	29.941	0.695	0.426
(b) Locomotion					
pointer	6906.361	1	6906.361	15.816	< .001*
difficulty	32328.758	1	32328.758	19.590	< .001*
pointer * difficulty	469.447	1	469.447	0.887	0.358
(c) Implicit references					
pointer	0.140	1	0.140	7.031	0.026
difficulty	3.560	1	3.560	80.807	< .001*
pointer * difficulty	2.984e-4	1	2.984e-4	0.010	0.924

* p < .005

5.2 Locomotion

We statistically compare the measures of locomotion (i.e. time spent moving) by performing a two way repeated measure ANOVA (Table I and Figure 5a). The ANOVA analysis results show two main effect related to the factors *Pointer* (p-value <.001) and *Difficulty* (p-value <.001). While we see an effect of locomotion related to the differences in the task, the important result is the effect on the pointer level and the lack of interaction between the two levels. When participants do not have a laser pointer, they spend approximately 18s more moving. To give a contextual understanding of this measure, the average duration of a task is 230 seconds, therefore representing approximately 7% of the time. If we consider that the average locomotion speed for this experiment is set to 1.6 m/s. This means that participants without support for distance pointing travelled approximately 28 meters more (in a 3m x 3m visualization space).

5.3 Implicit References

We statistically compare the repeated measures of the dependent variable: *number of implicit Deixes* by performing a two way repeated measure ANOVA (Table I and Figure 5a). The ANOVA analysis results show two main effects related to the factor: *difficulty* (p-value <.001). This result validates the design level of difficulty: if the referent is simple to identify by an explicit reference, the user tends to verbally describe it. On the other hand, when the referent is difficult to identify by verbal description, the user will adopt the strategy of pointing it and adding implicit references.

6 Discussion

Previous studies based on distance pointing in ICVE and real-world scenarios show that collaborators pointing accuracy from a distance often depends on either having access to a laser pointer or on how hard to describe it the referent (Wong and Gutwin, 2010, 2014). However, ICVE allows participants to move in the environment and, therefore, get as close as they need to the referent to perform an accurate pointing gesture. Therefore, what would users do when faced with the option of moving closer to the referent or describing it in better detail? Such a question is worth answering to understand better the dynamics of pointing-based communication in ICVEs. A better understanding of such collaborative dynamics is fundamental to developing solutions that can better support collaboration in ICVEs. Therefore within this study, we introduce the ability for users to move in the ICVE to investigate the trade-off between moving close to a referent and the effort of composing a verbal reference when the referent is difficult to describe. We do so within the context of a collaborative visual search task which is recognised to be a proxy of many other collaborative tasks in VR Prilla (2019).

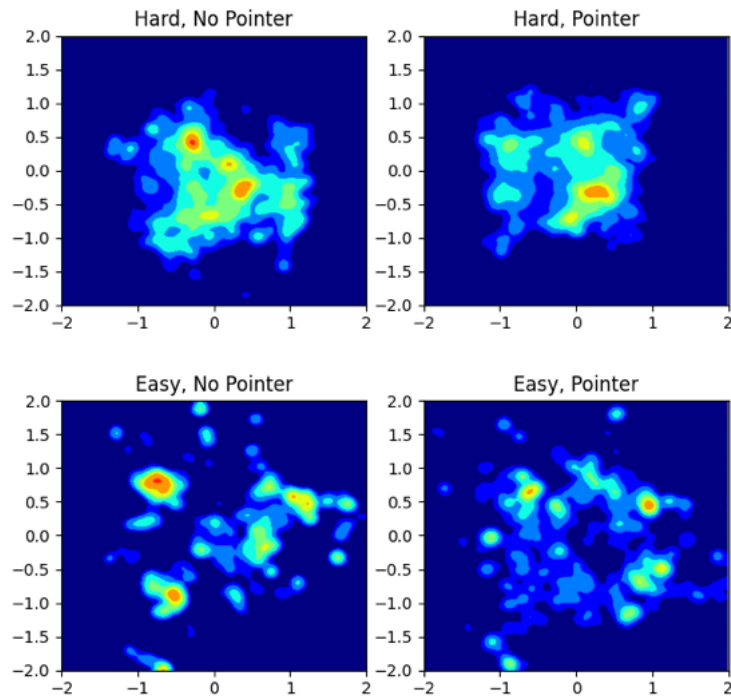


Figure 6: Heat-map of physical movement for the 4 experimental conditions.

6.1 Impact of locomotion on pointing-based communication

Our results extends the work of Wong and Gutwin (2010, 2014) by exploring a different dynamic of pointing based communication in the collaborative search task. While Wong measured accuracy in the context of fixed user distances from the referent, we explore a more ecologically valid scenario. Users are free to move in the ICVE and are instructed to perform a generalise search task. We extend his work by showing how users choose to locomote no matter how hard-to-describe is the referent in front of the choice of verbally describing a referent or moving closer to it. Such a statement is supported by the statistical analysis of locomotion, which shows a significant movement increment in hard and easy tasks when the pointer is absent.

Furthermore, we integrate the analysis of locomotion by generating cumulative head position heat maps for each experimental condition Figure 6. It is evident that the different datasets led to different exploration patterns and that the support for distance pointing did not impact how users explored the environment. If we cross the data from Figure 6 and Figure 5b and we notice that the locomotion last 20 seconds more in the absence of the pointer condition, we infer that such difference is not due to the exploration but to compensate lack of a laser pointer.

6.2 Impact of pointers on verbal communication

Previous CSCW studies in 2D desktop collaboration in remote programming show how pointers can increase the number of implicit references during deixis, making verbal communication faster and smoother (D'Angelo and Begel, 2017). Inspired by such a study, we counted and analysed the number of implicit references. In our ICVE experiment, results and observations suggest that when a pointer is not available, the number of implicit references (Fig 5c) during deixis stay the same. Our results differ from D'Angelo and Begel (2017) suggesting that when the embodiment is available, and users are free to move throughout the data pointers, visualisations do not influence verbal communication.

6.3 Impact of pointers on visual coordination

Previous research explored visual attention cues from head behaviour or eye gaze behaviour in ICVE during visual search tasks Piumsomboon et al. (2017) measuring how visual attention cues increase visual coordination. In general, hand pointing is recognized to trigger mutual orientation and visual coordination Wong and Gutwin (2010); Moore et al. (2007), however to the best of our knowledge, no study measure visual coordination with and without laser pointers in ICVEs. Our study fills this gap by showing that hand pointers availability increases the amount of time that collaborators spend concurrently pointing their heads towards the same subset of the data (section 5a).

7 Future work and Design Implications

In this study, we answered the following question: what will users do when faced with a lack of pointing accuracy: moving closer to the referent or describing it in better detail? While pointers in VR are proved extremely useful from previous studies Hindmarsh et al. (1998); Hoppe et al. (2018); Bai et al. (2020), we observe that visual pointers inclusion might depend on several factors: the complexity of the user interface, how crowded the ICVE is, and the confusion that multiple pointers may cause. Such considerations impact the design of ICVE, which needs to balance the advantages and disadvantages of pointers, compensating with alternative approaches that help to point accuracy. In addition, since there are benefits in moving closer to a referent, such as observing it in more detail or improving engagement with collaborators, we aim to identify methods that allow participants to semi-automatically move closer to an intended referent with or without pointing at it. A further approach can be identifying the intended referent by leveraging shared focus or adding semantic augmentation.

Our study does not consider distance perception as a crucial factor. This assumption is inherited from different works Mayer et al. (2020, 2018, 2015); Schweigert et al. (2019); Sousa et al. (2019); Wong and Gutwin (2014) that conversely consider distance with an active role in pointing accuracy. However,

this possible implication of distance perception in deictic pointing could be a good topic for future studies, as the research community is not yet detailed; studies that explore the perception of distance in VR are Finnegan et al. (2016); Maruhn et al. (2019).

Another interesting aspect is the implication of different locomotion strategies in ICVEs. For example, teleportation is a locomotion method which requires pointing to translate a user's location in the ICVE. Such a technique depends on the individual and the environment. However, our study, which explores the relations between pointing and locomotion, could inspire the community to investigate a collaborative version of locomotion. For example, when someone is making a pointing reference, the system can offer a "privileged" position and orientation for the observer that can be instantly applied. In addition, such a mechanism can be used for different collaboration tasks.

Moreover, we hope that the research community could use our results to explore novel ways of referencing targets based on a different paradigm or input channels such as speech. Previous studies demonstrate that a natural language processing pipeline could be used to describe and possibly display visual cues on some specific object parts Giunchi et al. (2021). Our study entails that when the referent is easy-to-describe, such a speech-based system could be used to highlight referents, such as collaborators are doing this naturally during a collaboration task. On the other hand, if the referent is hard-to-describe, that system may not be effectively used.

8 Conclusions

This paper designed and carried out an experiment to test the participants' attitude in a pointing-based task in ICVE. We conclude that deictic referencing in ICVEs with embodiment and locomotion does not require pointers to be accurate and implicit, as long as the users are free to move as close as they need to the data they are observing. One main reason is that when users are facing the problem of inaccuracy during pointing, they instinctively move closer to the referent rather than using verbal references to improve the precision of their pointing. Moreover, this effect is independent of how hard-to-describe the referent is. Locomotion allows users to move closer to the referent while performing deixis, improving pointing accuracy. We outline some design implications by highlighting how designers and engineers should consider two essential elements in support of distance-pointing: first, if users are able to move within the environment, and second if the collaborative task requires high visual coordination.

References

Bai, H., P. Sasikumar, J. Yang, and M. Billinghurst (2020): 'A User Study on Mixed Reality Remote Collaboration with Eye Gaze and Hand Gesture Sharing'. In: *Conference on Human Factors in Computing Systems - Proceedings*.

- Bangerter, A. and D. M. Oppenheimer (2006): 'Accuracy in detecting referents of pointing gestures unaccompanied by language'. *Gesture*, vol. 6, no. 1, pp. 85–102.
- Batmaz, A. U. and W. Stuerzlinger (2019): 'Effects of 3d rotational jitter and selection methods on 3d pointing tasks'. In: *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings*. pp. 1687–1692.
- Benford, S., J. Bowers, L. E. Fahlen, C. Greenhalgh, and D. Snowdon (1995): 'User embodiment in collaborative virtual environments'. *Conference on Human Factors in Computing Systems - Proceedings*, vol. 1, pp. 242–249.
- Biguer, B., M. Jeannerod, and C. Prablanc (1982): 'The coordination of eye, head, and arm movements during reaching at a single visual target'. *Experimental Brain Research*, vol. 46, no. 2, pp. 301–304.
- D'Angelo, S. and A. Begel (2017): 'Improving communication between pair programmers using shared gaze awareness'. *Conference on Human Factors in Computing Systems - Proceedings*, vol. 2017-Janua, pp. 6245–6255.
- Finnegan, D. J., E. O'Neill, and M. J. Proulx (2016): 'Compensating for distance compression in audiovisual virtual environments using incongruence'. In: *Conference on Human Factors in Computing Systems - Proceedings*. pp. 200–212.
- Giunchi, D., A. Sztrajman, S. James, and A. Steed (2021): 'Mixing modalities of 3D sketching and speech for interactive model retrieval in virtual reality'. In: *IMX 2021 - Proceedings of the 2021 ACM International Conference on Interactive Media Experiences*. pp. 144–155, Association for Computing Machinery.
- Higuch, K., R. Yonetani, and Y. Sato (2016): 'Can eye help you?: Effects of visualizing eye fixations on remote collaboration scenarios for physical tasks'. In: *Conference on Human Factors in Computing Systems - Proceedings*. New York, NY, USA, pp. 5180–5190, Association for Computing Machinery.
- Hindmarsh, J., M. Fraser, C. Heath, S. Benford, and C. Greenhalgh (1998): 'Fragmented interaction: Establishing mutual orientation in virtual environments'. In: *Proceedings of the ACM Conference on Computer Supported Cooperative Work*. New York, New York, USA, pp. 217–226, ACM Press.
- Hoppe, A. H., K. Westerkamp, S. Maier, F. van de Camp, and R. Stiefelhagen (2018): 'Multi-user collaboration on complex data in virtual and augmented reality'. *Communications in Computer and Information Science*, vol. 851, pp. 258–265.
- Jermann, P. ., D. . Mullins, and M.-A. . Nüssli (2011): 'Collaborative Gaze Footprints Correlates of Interaction Quality'.
- Kim, S., G. Lee, N. Sakata, and M. Billinghurst (2014): 'Improving co-presence with augmented visual communication cues for sharing experience through video conference'. In: *ISMAR 2014 - IEEE International Symposium on Mixed and Augmented Reality - Science and Technology 2014, Proceedings*. pp. 83–92, IEEE.
- Liu, C., O. Chapuis, M. Beaudouin-Lafon, and E. Lecolinet (2017): 'CoReach: Cooperative gestures for data manipulation on wall-sized displays'. *Conference on Human Factors in Computing Systems - Proceedings*, vol. 2017-May, pp. 6730–6741.
- Maruhn, P., S. Schneider, and K. Bengler (2019): 'Measuring egocentric distance perception in virtual reality: Influence of methodologies, locomotion and translation gains'. *PLoS ONE*, vol. 14, no. 10.

- Mayer, S., J. Reinhardt, R. Schweigert, B. Jelke, V. Schwind, K. Wolf, and N. Henze (2020): 'Improving Humans' Ability to Interpret Deictic Gestures in Virtual Reality'. In: *Conference on Human Factors in Computing Systems - Proceedings*. pp. 1–14.
- Mayer, S., V. Schwind, R. Schweigert, and N. Henze (2018): 'The effect of offset correction and cursor on mid-air Pointing in real and virtual environments'. *Conference on Human Factors in Computing Systems - Proceedings*, vol. 2018-April, pp. 1–13.
- Mayer, S., K. Wolf, S. Schneegass, and N. Henze (2015): 'Modeling distant pointing for compensating systematic displacements'. In: *Conference on Human Factors in Computing Systems - Proceedings*, Vol. 2015-April. pp. 4165–4168.
- Moore, R. J., N. Ducheneaut, and E. Nickell (2007): 'Doing virtually nothing: Awareness and accountability in massively multiplayer online worlds'. *Computer Supported Cooperative Work*, vol. 16, no. 3, pp. 265–305.
- Nüssli, M.-A. (2011): 'Dual Eye-Tracking Methods for the Study of Remote Collaborative Problem Solving'. *PhD Thesis, ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE*, vol. 5232.
- Pelz, J., M. Hayhoe, and R. Loeber (2001): 'The coordination of eye, head, and hand movements in a natural task'. *Experimental Brain Research*, vol. 139, no. 3, pp. 266–277.
- Pfeiffer, T., M. E. Latoschik, and I. Wachsmuth (2008): 'Conversational pointing gestures for virtual reality interaction: Implications from an empirical study'. In: *Proceedings - IEEE Virtual Reality*. pp. 281–282.
- Pietinen, S., R. Bednarik, T. Glotova, V. Tenhunen, and M. Tukiainen (2008): *A Method to Study Visual Attention Aspects of Collaboration: Eye-Tracking Pair Programmers Simultaneously*.
- Piumsomboon, T., A. Dey, B. Ens, G. Lee, and M. Billinghurst (2017): 'CoVAR: Mixed-Platform Remote Collaborative Augmented and Virtual Realities System with Shared Collaboration Cues'. In: *Adjunct Proceedings of the 2017 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2017*. pp. 218–219, Institute of Electrical and Electronics Engineers Inc.
- Prilla, M. (2019): '"I simply watched where she was looking at": Coordination in short-term synchronous cooperative mixed reality'. *Proceedings of the ACM on Human-Computer Interaction*, vol. 3, no. GROUP.
- Šašinka, C., Z. Stachoň, M. Sedlák, J. Chmelík, L. Herman, P. Kubíček, A. Šašinková, M. Doležal, H. Tejkl, T. Urbánek, H. Svatoňová, P. Ugwitz, and V. Juřík (2019): 'Collaborative immersive virtual environments for education in geography'. *ISPRS International Journal of Geo-Information*, vol. 8, no. 1.
- Schmalstieg, D. and T. Höllerer (2016): *AR Textbook Tobias*.
- Schmidt, C. L. (1999): 'Adult Understanding of Spontaneous Attention-Directing Events: What Does Gesture Contribute?'. In: *Ecological Psychology*, Vol. 11. pp. 139–174.
- Schneider, B. and R. Pea (2013): 'Real-time mutual gaze perception enhances collaborative learning and collaboration quality'. *International Journal of Computer-Supported Collaborative Learning*, vol. 8, no. 4, pp. 375–397.
- Schroeder, R., A. Steed, A. S. Axelsson, I. Heldal, A. Abelin, J. Widstrom, A. Nilsson, and M. Slater (2001): 'Collaborating in networked immersive spaces: as good as being there together?'. *Computers & Graphics*, vol. 25, no. 5, pp. 781–788.

- Schweigert, R., V. Schwind, and S. Mayer (2019): 'EyePointing: A Gaze-Based Selection Technique'. vol. 19.
- Slater, M., A. Sadagic, M. Usoh, and R. Schroeder (2000): 'Small-group behavior in a virtual and real environment: A comparative study'. *Presence: Teleoperators and Virtual Environments*, vol. 9, no. 1, pp. 37–51.
- Sousa, M., R. K. Dos Anjos, D. Mendes, M. Billinghurst, and J. Jorge (2019): 'Warping deixis: Distorting gestures to enhance collaboration'. In: *Conference on Human Factors in Computing Systems - Proceedings*, Vol. 12. New York, NY, USA, pp. 1–12, Association for Computing Machinery.
- Step toe, W., O. Oyekoya, A. Murgia, R. Wolff, J. Rae, E. Guimaraes, D. Roberts, and A. Steed (2009): 'Eye tracking for avatar eye gaze control during Object-Focused multiparty interaction in immersive collaborative virtual environments'. In: *Proceedings - IEEE Virtual Reality*. pp. 83–90, IEEE.
- Villamor, M. and M. M. Rodrigo (2018): 'Predicting successful collaboration in a pair programming eye tracking experiment'. In: *UMAP 2018 - Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization*. pp. 263–268.
- Wang, P., S. Zhang, X. Bai, M. Billinghurst, W. He, S. Wang, X. Zhang, J. Du, and Y. Chen (2019): 'Head pointer or eye gaze: Which helps more in MR remote collaboration'. In: *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings*. pp. 1219–1220, Institute of Electrical and Electronics Engineers Inc.
- Widestrom, J., A. S. Axelsson, R. Schroeder, A. Nilsson, I. Heldal, and A. Abelin (2000): 'The Collaborative Cube Puzzle: A Comparison of Virtual and Real Environments'.
- Wong, N. and C. Gutwin (2010): 'Where are you pointing? The accuracy of deictic pointing in CVEs'. In: *Conference on Human Factors in Computing Systems - Proceedings*, Vol. 2. New York, New York, USA, pp. 1029–1038, ACM Press.
- Wong, N. and C. Gutwin (2014): 'Support for deictic pointing in CVEs'. pp. 1377–1387.

Bardou Maxime, Letondal Catherine, Imbert Jean-Paul, Causse Mickaël, Bidegaimberry Maxence, Dubus Romane, Marcon Cécile (2022): Redesigning systems for Single-Pilot Operations: the mutual awareness problem for remote crews. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_ep03

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_ep03

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

Redesigning systems for Single-Pilot Operations: the mutual awareness problem for remote crews

Bardou Maxime, Letondal Catherine, Imbert Jean-Paul, Causse Mickaël, Bidegaimberry Maxence, Dubus Romane, Marcon Cécile
Interactive Informatics, University of Toulouse - ENAC, Toulouse, France
Contact Author: maxime.bardou@enac.fr; catherine.letondal@enac.fr; jean-paul.imbert@enac.fr; mickael.causse@isae.fr

Abstract. Currently, flight safety is ensured by the collaboration of at least two pilots in the cockpit. Thanks to progress in automation and telecommunication, aircraft manufacturers and aviation companies envision that a single pilot in the cockpit assisted by a pilot on the ground (i.e Single-Pilot Operation) could ensure flight operations while requiring less human resources. However, without appropriate collaboration tools, this situation of remote collaboration may lead to a degradation of the awareness of actions and attitudes between the two pilots (i.e. mutual awareness). In this paper, we propose to enrich the understanding of the remote collaboration problems of two pilots through a fine-grained analysis of mutual awareness needs. First, we describe awareness frameworks from the literature. Second, we identify awareness issues during a case study involving a crew of pilots in two distant flight simulators. Third, we refine the relevant awareness concepts through exploratory prototyping of collaborative tools. These prototypes are based on three scenarios involving specific awareness requirements including 1) visualizing the physiological state of the pilot on board during a non stabilized approach, 2) an emergency decision making, and 3) global awareness during a whole flight for a better efficiency of the ground assistant operator at the arrival. In this article our contribution is a

refined study of the awareness needs adapted to the context of remote collaborative piloting, with the final objective of designing more appropriate tools.

1 Introduction

In the coming years, the pilot's task will evolve with the arrival of a new, more automated single-pilot cockpit (Council et al., 2014). The emergence of Single-Pilot Operation (SiPO) bring new operational issues for aviation safety. To make this concept as safe as the two-pilot configuration, several scenarios are foreseen. One of them is the assistance of the pilot on board the cockpit by a ground pilot. However, this remote collaboration between on-board and ground pilot generates new problems such as the loss of awareness of the other pilot's actions and attitudes, namely the mutual awareness. The concept of SiPO is not just a question of separation between the pilot and co-pilot. This new concept leads to a modification in their collaboration with new tasks, new tools and new automatisms. Tools have already been proposed by different authors (Lachter et al., 2017) to overcome the distance issue. However, the problem would need to be better conceptualized to ensure that the tools are well fitted with remote collaboration.

Concepts of remote collaboration tools to support awareness between collaborators are already widely studied in the Computer-Supported Cooperative Work (CSCW) literature (e.g. (Beaudouin-Lafon and Karsenty, 1992; Gross, 2013; Greenberg and Gutwin, 2016; Schmidt, 2002)). Nevertheless, awareness needs between a pilot and a ground pilot during remote collaboration are not yet explicit. This article aims to answer several questions : what awareness framework are applicable in a new aeronautical context? What are the awareness needs in remote collaboration in a Single-Pilot operation context? And finally how awareness can be integrated into future collaborations tools?

To answer these questions, we will set the context, explore the challenges and hypothetical solutions to the concept of SiPO existing in the literature. In addition, we use the Antunes et al. (2014) concept of awareness and the mutual awareness concept of Schmidt (1998) in the context of the piloting activity and explore the concepts associated with collaboration tasks, which are coordination, communication and cooperation. To support our research, we also present the results from interviews that we conducted with three pairs of pilots in situations of remote collaboration in flight simulators (section 5). Finally, we describe how exploratory prototyping helped us to refine concepts of awareness in SiPO context (section 6). This exploration enabled to reflect on design in three scenarios involving specific needs of awareness. We identified a need for the ground assistant to visualize the physiological state of the pilot on board during a non stabilized approach. Then, we explored shared map for an emergency involving a rerouting. Finally, we studied how to improve the global awareness during a flight

for a better efficiency of the ground assistant operator at the arrival. We conclude this paper by a discussion of our contribution and of future works.

2 Context

In the 1950s, the flight deck was composed of five members, the pilot, the co-pilot, the radio operator, the navigator, and the mechanic (Bohn, 2010). In the 1980s, the evolution of technology and the development of new on-board systems allowed the size of the crew to be drastically reduced from five to two members (pilot and co-pilot). This technological and operational evolution required new tools, new cockpit concepts (e.g., glass cockpit), as well as new safety procedure such as the Crew Resource Management (CRM) (a set of procedures to train the crew to reduce human errors). Currently, airlines companies are looking to reduce the number of crew member in a cockpit to reduce their costs. To make this possible, several scenarios are foreseen.

Over the past decade, a new phase of crew reduction in commercial aviation began to be studied by the aeronautical industry. This new phase includes two successive concepts of operations. Firstly, the extended Minimum Crew Operation (eMCO) or Reduced Crew Operation (RCO). And secondly, the Single-Pilot Operation (SiPO). The eMCO allowed to reduce the crew from 2 to 1, only during the cruise phases (e.g Connect project (Airbus, 2022)). In the case of SiPO, this reduction would be effective along the whole flight (e.g. DISCO project (Airbus, 2022)(Fig.1)).

According to European Union Aviation Safety Agency (2021) (EASA), the challenges for eMCO and SiPO in terms of safety are numerous: pilot error, monitoring pilot performance, pilot incapacitation, etc. To meet these challenges and achieve a level of safety equivalent to today's two-pilot operations, solutions must be found, such as ground pilot, virtual assistant, advanced cockpit design with workload alleviation means, capability to cope with pilot incapacitation etc..



Figure 1. DISruptive COckpit project (DISCO) by Airbus that would allow the eMCO and SiPO concept.

2.1 Ground assistant in SiPO context

In this section, we present the consequences of the SiPO concepts and the various proposals made by researchers to maintain flight safety.

Schmid and Stanton (2020) describe various problem in the concept of SiPO: operational, automation, pilot incapacitation, communications, and certifications problem. Some of these problems, such as incapacitation and certification, are not the focus of this paper but have been studied by many authors (Paz Goncalves Martins et al., 2021; Schmid and Stanton, 2019). Some projects such as HARVIS (Human Aircraft Roadmap For Virtual Intelligent System) (Duchevet et al., 2020), NiCO (Next generation Intelligent Cockpit) (Niermann and Kügler, 2021), and the Clean Sky project (CleanAviation, 2022), focused on a virtual assistant to assist the pilot on board in specific situations where the assistance of a third party is required. For example, in Duchevet et al. (2020), the virtual assistant provides help for decision making during non-stabilized approaches by alerting the pilot about a deviation from his trajectory.

Other authors are working on the creation of a new cockpit and ground station to assist the pilot on board (Niermann and Kügler, 2021; Lachter et al., 2014b). All these proposals implies different technologies. Some solutions are oriented towards human-machine interaction while other are oriented towards human-human interaction.

2.1.1 Operational solutions for ground assistance

Proposals by industry and researchers for ground support raises new problems of organization and roles between the two pilots. To do so, three concept of operations have been made to support the on-board pilot thanks to a ground assistance: the Hybrid Ground Operator Unit (HGO), the specialist Ground Operator Unit (SGO), and the Harbor Pilot concept (Bilimoria et al., 2014). These three proposals are distinguished by the allocated tasks to the ground assistants (see Fig. 2).

The Hybrid Ground Operator Unit will carry out three types of activities:

- conventional dispatcher tasks (management of flight plans and routes, communication with airlines etc.)
- the monitoring of multiple nominal flights
- the individual support of non-nominal flights

The hybrid ground assistant is dedicated solely to the non-nominal flight in order to provide exclusive support to the aircraft concerned.

The second proposal by Bilimoria et al. (2014) is the Specialist Ground Operator Unit. In this case, two groups form the unit: the Ground Associate and the Ground Pilot or Dedicated Pilot. The first group will be in charge of dispatch duties and nominal flight monitoring. Contrary to the Hybrid Ground Operator unit, ground associates will not deal with non-nominal flights. These flights will be redirected to the dedicated pilot who will be responsible for the individual support of the aircraft in critical situation. When the dedicated pilot does not support the

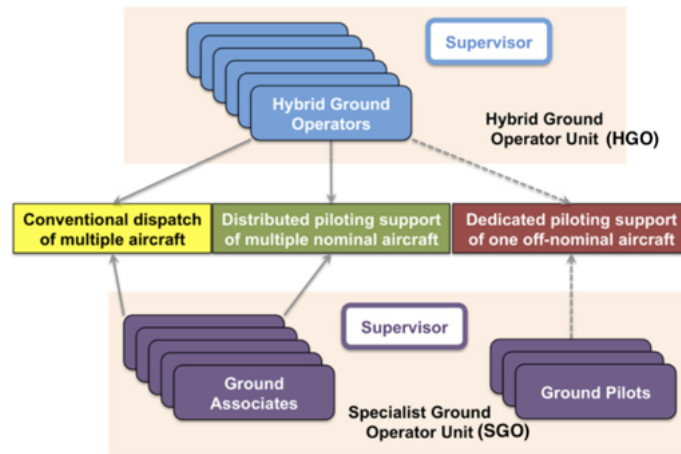


Figure 2. The two ground unit possibilities (Hybrid and Specialist Ground Operator) with in the middle the three tasks of ground unit (Bilimoria et al., 2014).

non-nominal flights, the dedicated pilot will be "on standby or performing collateral duties" (Bilimoria et al., 2014).

In addition to the other two proposals, the Harbor Pilot will support the aircraft in a defined area known to be complex in the case of aircraft ascending or descending. The objective is to decrease the workload of the pilot on-board.

3 Related Work

In this related work, we will present studies about the impact of remote collaboration in SiPO context and the authors' approach to the problem of remote pilot collaboration. In a second step, we will try to address the problem of distance between pilots with an approach focused on awareness during remote collaboration.

3.1 Impact of remote collaboration in allocated pilots tasks

Today, the pilot's tasks are divided into four categories: piloting, navigating, communicating and systems management (Billings, 1997). Each task requires cooperation, coordination and communication between pilots. However, these collaboration processes are achieved through the co-location of the pilots inside the cockpit. With the change of cockpit (automation, crew member, system) and the existence of a ground pilot, the collaboration between pilots will be impacted by the distance. Indeed, the physical separation of two pilots in the actual cockpit cause the loss of non-verbal communication (such as gestures, posture, etc.). To be more precise, the loss of non-verbal communication leads to uncertainty of roles between pilots, uncertainty of actions, uncertainty of manipulation, gathering information and decision making problem (Lachter et al., 2014a). This loss of

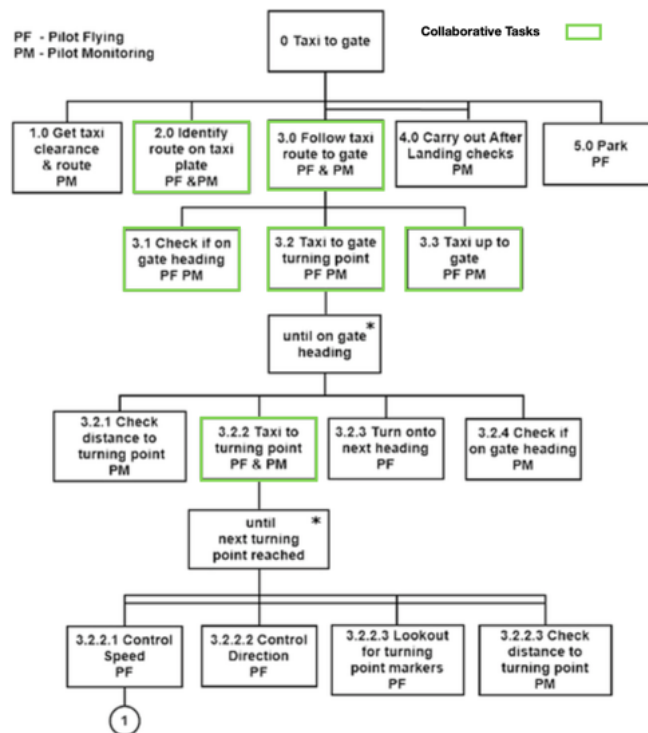


Figure 3. Hierarchical task analysis during taxi to gate with collaborative tasks in green adapted from Huddleston et al. (2015).

non-verbal communication will therefore disrupt the coordination and consequently the collaboration between the two pilots. For example, the "taxi to the gates" phase requires five collaborative tasks between the two pilots. If the two pilots do not coordinate for the identification of the taxiway (see Fig. 3 : 2.0 "Identify route on taxi Plate"), either through the physical environment (direct view of the runways) or the airport layout, then they could, for example, miss the turn off. In such situation, the integration of new collaboration tools are needed to avoid a degradation of the collaboration.

The studies of remote collaboration in SiPO context (Lachter et al., 2014a,b; Ligda et al., 2015; Brandt et al., 2015; Lachter et al., 2017) show us a lack of awareness in remote collaboration that is addressed by several types of tools. In Lachter et al. (2017), three prototypes of ground station with collaboration tools have been designed and tested (see Fig. 4).

In this approach, the aim is to address the uncertainties in action, manipulation, role, the problem of gathering information and decision making (see section 3.1). For the uncertainty of the roles between the "Pilot Flying" (PF) and "Pilot Non Flying" (PNF) (currently called "Pilot Monitoring" (PM)), the addition of a Crew Resource Management indicator was implemented in the cockpit on board and in the Ground Station. This tool, consisting of 6 LCD touch screens, allows the tracking of responsibilities and actions. One of the screens allows the allocation of roles by displaying "PF" or "PNF". When the role is assigned to one of the pilot,

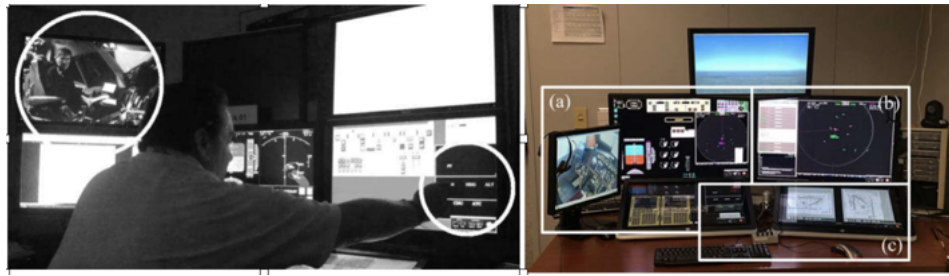


Figure 4. Ground Station (Lachter et al., 2014b) (left). The SiPO III ground station (Lachter et al., 2017; Dao et al., 2015) (right) (a) replicated flight deck displays for the chosen aircraft; (b) flight tracking displays with ELP recommendations; and (c) crew collaboration tools including shareable charts; (d) the Traffic Situation Display and Emergency Landing Planner recommendations (lower left corner).

the other five screens associated with the pilot's tools turn green if the pilot is "PF", or white if the pilot is "PNF". For example, if the first screen showed "PF", then LCD screens below the Mode Control Panel which show "SPD", "HDG", "ALT" turn green and the screen which show "ATC" and "CDU" turn white (see Fig. 4). The CRM indicator also solves the problem of uncertainty of manipulations and actions by indicating with an arrow if the changes in speed are going up or down (manipulations). When the desired value is reached, the speed is displayed and must be validated by the pilot by pressing the touch screen. In addition to the CRM indicator, the video feedback made it possible to see the actions of the two pilots and what they were watching (e.g. MCP, CDU, checklist, etc.). Finally, for gathering information and decision making, the shared maps contained in the Electronic Flight Bag (EFB) could be consulted in a synchronized way.

3.2 Awareness approaches and concepts

Awareness in collaboration tools is an extensive topic with many proposals (Gutwin et al., 1996; Hill and Gutwin, 2003; Greenberg and Gutwin, 2016; Bravo et al., 2013). During collaboration the pilots must have good situation awareness (the perception, comprehension and projection of action and event in our own environment) (Endsley, 2017). To obtain this situation awareness, the pilots must be aware of each other's implicit and explicit activities during the remote collaboration (i.e mutual awareness).

In order to specify whether the integration of awareness is taken into account, Antunes et al. (2014) has created a more complete checklist taking into account a bigger number of awareness concepts than Gutwin et al. (1996) and a questionnaire making it possible to evaluate awareness in groupware. Antunes et al. (2014) showed that six awareness concepts can be used for collaboration tools: collaboration awareness, location awareness, context awareness, situation awareness, social awareness and workspace awareness. This framework of awareness will be used to interpret the results of our case study.

3.2.1 Situation Awareness

Situation awareness allows a "generalization of the notion of workspace". Indeed, the working environment (cockpit, ground station) plays a role in the decisions of both pilots. In order to obtain adequate decisions, the pilots must perceive their environment, understand it and project the consequences of their actions on the environment (Endsley, 2017). In the case of remote collaboration in SiPO, the two pilots do not have the same environment, so it is crucial that each pilot is aware of the impact of their action on the collaborator's environment.

3.2.2 Mutual awareness: a consideration of implicit actions

In a cockpit, pilots will manage to coherently integrate their actions in a discrete way in order to achieve a common goal. This awareness of each other's activity can be communicated explicitly (oral or digital communication) or implicitly (intonation in the voice, modulation of gestures) during an action. These are the implicit actions, called "discrete action" (Schmidt, 1998).

According to this author, the awareness between two people is not only an explicit fact allowing to coordinate these activities. It is a whole that also includes the implicitness of actions. This set called mutual awareness is "the perception and understanding of the activities of member A, including the intention, status and possible outcomes of that activity, by member B" (vice versa) (Schmidt, 1998). In a remote collaboration, it is a question of a loss of information that the collaborator transmits, what it looks like, whether it is permanent to whom it is transmitted and what the intentions of those who use this information are going to be (Bellotti and Sellen, 1993).

Mutual awareness has already experienced problems in the field of aeronautics and more precisely in touch screens (Becquet et al., 2019). Indeed, the loss of tangible elements leads to an abstraction of actions (implicit or explicit) and a loss of information for the collaborator during the manipulation of an object. More generally, mutual awareness emphasizes the combination of attention directed towards a collaborator (focus) and the way in which this collaborator projects his presence and his activities towards him (nimbus).

3.2.3 Collaboration Awareness

Collaboration awareness corresponds to the perception of the availability of the two collaborators (Group availability) and the mode of communication used. The availability of the group of pilots is distinguished by the relative position of the two collaborators (same place or different place) (Johansen et al., 1991) and if the pilots are online or not (Schmidt, 2002). Concerning the mode of communication, it can be synchronous, asynchronous or even semi-synchronous. This reduction in mental distance will allow a better projection of the situation for the ground support and the feeling of being in a crew. The characteristics of this awareness allow us to suppose its importance in communication tasks, especially for the need to obtain

the availability of the collaborator but also for the awareness of the communication mode.

3.2.4 Location Awareness

According to Antunes et al. (2014), location awareness refers to the geographical position of collaborators and more precisely to the awareness of the position, topology and attributes of the space (e.g. weather, temperature). In the case of a non-nominal situation, it is important for the ground support to know the geographical position of the pilot on board. Indeed, the indication of weather conditions (storm, wind) and also topography (sea, land, mountain) allows ground pilot to adapt his decision making.

3.2.5 Context Awareness

Context awareness is related to virtual space (computer-based interactive spaces) (Rodden, 1996). It allows the understanding of changes and events in the virtual space to be maintained between two collaborators. In the case of piloting, the pilots will know where the collaborator is in the virtual workspace like an electronic checklist. For example if the pilot is on page "X", the second pilot will not have to ask where he is in the checklist.

3.2.6 Social Awareness

Social awareness corresponds to the social context during the collaboration between the two pilots. The social context is the understanding between the two pilots of each other's roles and activities but also of what and how the members of the group contribute to a task (Carroll et al., 2003). Indeed, within the framework of the SiPO concept, the roles and activities must be clear when collaborating at a distance. As noted by Lachter et al. (2014a), confusion has been observed regarding the roles assigned to the pilots and their activities while their definition is important in piloting activity.

3.2.7 Workspace awareness

According to Snowdon et al. (2000), the workspace corresponds to "a container of places with continuous activities". It is possible to focus on place, which allows the organization of tasks (who, what, when, how) as allowed by group editors (writing, revising, global activity view) (Koch and Koch, 2000). For example, the on-board pilot knows that the pilot on the ground (red tack, who) is in the process (real time, when) of highlighting (tack, what) an airport on a shared map, thanks to a marking system (planting tack, how).

3.2.8 Informal Awareness

We can identify an additional awareness possibly impacted by SiPO : the informal Awareness related to the need for informal communication. Indeed, according to Röcker (2012), informal awareness is "similar to informal communication". As far as informal communication is concerned, it is necessary for the smooth running of a crew because it allows the two pilots to obtain an awareness preview of the situation. This informal communication makes it possible to defuse situations before they become critical via discussions in the corridor or with one's office neighbor (Mackay, 1999). In the case of a standard configuration in the cockpit, the two pilots can have informal discussions during phases when the workload is reduced (cruise phase) and obtain a more important situation awareness.

Awareness	Example in a flight context	Problems identified by Lachter et al. (2014b)
Collaboration	The pilot on the ground knows that the pilot on board is available to collaborate	No problems identified
Location	Both pilots know where the other is geographically located	No problems identified
Situation	Both pilots see and understand each other's actions and their consequences on the cockpit	Action uncertainty
Social	What roles are assigned to each of the two pilots and what are their activities	Role uncertainty
Context	Both pilots know which page the pilot is on when consulting the QRH	Action uncertainty Gathering information Decision making
Workspaces	Both pilots know who, what, when and how tasks are performed in the workspace	Manipulation uncertainty Gathering information Decision making

Table I. Categorization of uncertainties identified by Lachter et al. (2014b) through Antunes' Awareness in a flight context.

In this paper, we choose to study the awareness proposed by Antunes et al. (2014) for his more comprehensive list of awareness as well as the informal awareness described by Röcker (2012). The application of the Antunes et al.

(2014) framework on the problems raised by Lachter et al. (2014a) allows us to bring some precisions on the lack of awareness in SiPO context (see Table I.).

4 Methodology

To analyze the awareness problem as accurately as possible, we conducted a series of activities. The first step was to study actual pilot task in real flight conditions and in simulations, using videos and interview transcripts generated in a previous project (Letondal et al., 2018). Following this study, we decided to set up a remote collaboration situation with the aim of clarifying awareness needs in collaboration with three pairs of commercial airlines pilots. In this case study, we separated two pilots in two different rooms in a low-fidelity cockpit (computer, flight simulator, and yoke) to observe the consequences of remote collaboration in a non-nominal flight (unusual situation). After this remote collaboration situation, we conducted an interview (1 hour, right after the simulation) with the pair of pilots to provide an analysis of awareness needs in remote collaboration. To extract the maximum information from these interviews, we identified the different insights associated with awareness and collaboration. Then, we associated different keywords to quotes such as "informal communication", "verbalization", "action", "perception", "action monitoring". Finally, we carried out exploratory prototyping to refine awareness concepts of SiPO.

5 Case study: a flight with the two pilots separated

In this section, we describe our case study that we conducted to understand the impact of separation. In this study, we chose, while using a current airliner cockpit (A320 cockpit simulator), to introduce an artifact by dividing the cockpit in two and putting each pilot in a different room. Our goal was thus to isolate the distance dimension by moving from co-located collaboration to remote collaboration without changing the rest. We were counting on this device to observe the effect of separation and distance in the behavior of the pilots, in particular to identify what in this behavior would be a problem of mutual awareness without perception of gestures, bodies, head directions, etc., of the other pilot. In addition, to accentuate the problems of mutual awareness, we opted for a non-nominal, but nevertheless usual situation in pilot training.

5.1 Participants and setting

In our experiment, six airline pilots were divided into three pairs of *Pilot Flying* and *Pilot Monitoring*. The 3 pairs of pilots can be categorized according to the number of flying hours (Table II.). Pairs A and B have a real experience of Multi-Crew Cooperation (MCC) experience, while Pair C has not MCC experience outside the simulator.

Couple	Age	Flight hours
P1.A	44	9200
P2.A	43	8400
P3.B	45	4500
P4.B	33	3500
P5.C	30	230
P6.C	21	410

Table II. Categorisation of participants.

5.2 Material

The pilots were separated in two different rooms. In each room, the pilots were placed in front of a computer running MICROSOFT FLIGHT SIMULATOR 2020 (MFS2020), with a community mod called YourControl, allowing cockpit sharing and synchronization of the two MFS2020 licences. To control the flight path, the pilots were equipped with a Honeycomb Alpha Flight Controls stick. The simulation takes place on an Airbus A320 and each pilot has a Quick Reference Handbook (QRH) associated with the aircraft.



Figure 5. Video editing of the two pilot in low fidelity cockpit and the cockpit screen capture.

5.3 Flight scenario

First of all, a briefing took place with the pilots concerning the flight conditions. After providing the flight briefing, the pilots were assigned the roles of pilot flying and pilot monitoring and went to two separate rooms. The flight, departing from Gran Canari (GCLP) to Gatwick (EGKK), started directly on the runway for the take-off phase. Pilots were only instructed that they were at FL350 during the cruise phase. On arrival at the waypoint (VASTO), a cabin smoke alert was announced by one of the fictitious cabin crew. Following this announcement, we decided to stop the simulation when the pilots arrived at the PACK 2 item of the

A320 QRH cabin smoke procedure. The whole scenario was recorded with two cameras facing each pilot and the screen recording of the MFS2020 simulator. Following the scenario, interviews were conducted for approximately one hour. These interviews were transcribed and analyzed qualitatively.

5.4 Results

As a result of the interviews and the transcriptions, we were able to gather some insights to specify needs to achieve mutual awareness. From the interviews, we identified the different awarenesses impacted by remote collaboration in a current cockpit.

Social Awareness needs During the remote collaboration, the coordination between all pairs of pilots was not equivalent. Some of them had coordination problems, as both pilots were not aware of the roles and activities that each of them had to perform at the beginning of the flight, while the roles were predefined during the briefing.

(1)(P6.C) : "so at the beginning we were a little bit [...] not well ordered because we didn't know who had to do what actions and all that [...]"

This confusion, which we can see in the quote (1), is caused by a lack of awareness regarding the distribution of roles. This finding is consistent with the problems identified by Lachter et al. (2014a) regarding role uncertainty during remote collaboration between the two pilots. Without the appropriate tools to make both pilots aware of each other's roles in the remote collaboration, we noticed that remote collaboration mainly impacted coordination.

Context Awareness needs Other problems identified during the interview concerned context awareness. Indeed, when handling the QRH, the two pilots had no idea of the position of the collaborator in the QRH when they consulted it as the following quote shows :

(2)(P6.C): "in fact it's not badly done in the sense that it has pages with numbers a moment ago I told you but wait [...] no, it was you who told me ... and so it was important to have a well-established QRH with references so that you don't get confused and it's also important to make it clear which page you're on, especially when you can't see yourself, especially when you can't read the page directly [...]"

This confusion can be easily solved by a computerized and synchronized QRH between the two pilots. This way both pilots will know where the pilot is looking

in the QRH. This need for awareness seems to be associated with coordination in the 3Cs model, in that elements of context awareness such as the other's view or position in the QRH allows them to improve their coordination and therefore their communication to reach a decision (cooperation).

Workspace awareness needs In our case study, verbal communication is accentuated to express actions that may not be perceptible to the second pilot. This emphasis allows to highlight some important elements to be transcribed at system level in the future SiPO cockpit. For example, the accentuation of FCU (Flight Control Unit) modifications on the ground pilot interface to allow on-board pilot to be aware the modifications without verbal communication.

It is important to note that the verbalization of one of the pilots make the other pilot, who is not aware of these characteristics, aware of the actions, intentions and emotional states. As a result, pilots do not know when, what, how, the action are made in their workspace (i.e the cockpit) without verbalization.

This increase in verbal communication can be associated with a lack of workspace awareness. As each pilot was not sure whether the other was aware of the other's actions and intentions, the pilot making the changes in the cockpit would verbalize everything he or she did. This is in contrast with current rules, as explained by this pilot :

(3)(P2.B) : "So we [...] we don't verbalize any more [...] in fact P3.B, being in manual piloting with change of altitude it's automatic that it's the PM who displays it and we have the non verbal action in the visual field the PF sees that there is an action on the switch of the altitude that there is [...] Whereas I was verbalizing it I put the altitude on and then I made the official announcement [...] the cross checker behind. Usually I don't say I'm putting the 350 level on automatically"

This accentuation of verbalization allow pilots to mentally reconstruct the actions performed in the cockpit as this quotation shows :

(4)(P3.B): "[...] I could see that he was setting things, I could see everything he was setting, everything he was setting he was saying so I could see it".

The accentuation of verbalization during our case study show us a lack of coordination which can be resolved through workspace awareness elements that show who, what, when and how actions are performed in the environment.

Informal awareness needs One of the operational problems of the SiPO concept concerns the feeling of crew and more precisely the relationship between two pilots in a cockpit.

(5)(P6.C) : "the fact that I knew him I found it easier to communicate with him [...] a little more fluid and I dared to speak more and propose things, whereas what I remember in MCC (Multi Crew Cooperation), the fact that I didn't know the person, I dared to propose things perhaps less"

(6)(P3.B) : "The fact that we were talking like that, you were next to me Yeah, so I had I didn't have the presence but I still had the presence of the voice"

Informal communication makes it possible to defuse situations before they become problematic (Mackay, 1999). This communication also contributes to the sense of crew, an important feeling among pilots. In the SiPO concept, the pilot alone cannot communicate informally with the ground if the ground is not available. This lack of awareness brings us very clearly closer to a communication problem in the 3Cs model.

5.4.1 Conclusion of the case study

From the different interviews we can see that most of the awareness concepts has been strongly impacted by remote collaboration in a standard cockpit. In the case of collaboration in a current cockpit: context awareness, social awareness, informal awareness and workspace awareness seem to be low. However, we observed that the location awareness (awareness of the position, topology and attributes of the space (e.g. weather, temperature) and collaboration awareness (perception of the availability of the two collaborators and the mode of communication used) described by Antunes et al. (2014) were intact because both pilots are in the same virtual cockpit synchronized by the add-on with direct communication. Concerning mutual awareness and situation awareness, we noticed that both awareness are impaired in remote collaboration in SiPO context. Indeed, situation and mutual awareness are part of an overall awareness when collaborating which allow pilots to coordinate, cooperate and communicate. Finally, to achieve this mutual awareness in SiPO context, a subset of awareness is needed as context awareness, social awareness, informal awareness and workspace awareness.

6 Exploratory Prototyping

Exploratory prototyping activities were conducted to refine our understanding of collaboration awareness, context awareness, workspace awareness, social awareness and situation awareness in situated and concrete scenarios involving an on-board pilot and a ground pilot. Table III provides an overview of explored awareness concepts in each prototype. Two pilots were involved during prototypes 2 to 7, one of whom participated in the simulation (P3.B) and a second pilot who was not present at any of the stages of our approach.

Awareness	Prototype number
Collaboration	1,2,3
Context	4,5
Situation	7
Social	6
Workspace	6

Table III. Table of the different prototypes proposed according to awareness.

The prototypes were designed on the basis of scenarios that are provided in Table IV. Two of the scenarios are based on non-nominal situation (Non-stabilized approach and Rerouting) and one is during a nominal situation (Taxiing).

Scenario	Description	Prototypes	Awareness
Non-stabilized approach	Non-stabilized approach of the aircraft requiring high physical and mental concentration.	3	collaboration
Rerouting	Rerouting solution due to a medical emergency requiring a decision between the pilot on board and the pilot on the ground.	1,2,4,5	context social workspace
Taxiing	Ground guidance scenario on arrival (taxiing) requiring tools to provide the ground operator assisting the pilot with a global awareness of the course of the flight.	6	situation

Table IV. Table of the different scenario according to prototypes.

6.1 Exploring collaboration awareness design

The aspect of collaboration awareness studied in the following prototypes are availability and communication mode as described by Antunes et al. (2014).

Availability has been explored in several forms: clock indicator, physiological data visualizations and by a "check mark" icon on a shared tool. The communication mode has been explored through a synchronous or asynchronous mode indicator in a shared map.

Prototypes 1 and 2 and their design elements presented below are integrated into a shared map for a rerouting scenario where one of the passengers suffers a heart attack. Faced with this medical situation, the pilot on board and the ground pilot

must agree on a diversion airport as soon as possible, according to different criteria and views (weather, terrain).

6.1.1 Prototype 1: a clock to indicate asynchronous actions

The first prototype is included in an asynchronous shared map between ground and on-board pilot. When the clock is displayed, it indicates that the other pilot is busy with an action in progress (e.g. a choice of diversion scenario) performed asynchronously (Fig. 6 and 7).

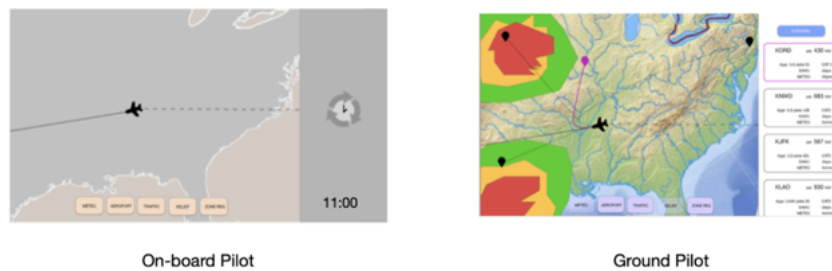


Figure 6. Prototype 1 Clock Design: Shared maps with on the left the on-board pilots' screen and on the right the ground pilot' screen. The ground pilot selects a diversion scenario on the left of his screen. On the left-hand side of the pilot's screen, a "clock" indicates that scenarios are being chosen by the pilot on the ground.



Figure 7. Prototype 1 Clock Design: Shared maps three minutes after choosing a diversion scenario, it is sent to the shared map of the pilot on board. The clock always present, indicates that a second scenario is chosen by the pilot on the ground.

6.1.2 Prototype 2: Availability check mark

A second availability indicator has been designed in the form of a green check mark when the other pilot is available and a red cross otherwise (Fig. 8). An overview of the other pilot's screen completes this information.

6.1.3 Prototype 3: Physiological state widgets

The 3rd prototype proposes indicators of the physiological state of the pilot on board. These indicators could for example be useful in an non-stabilized approach



Figure 8. Prototype 2: Shared maps with availability indicators on the left on on-board pilot' screen and ground pilot' screen. A) a green check-mark indicates the availability of the ground pilot to collaborate B) a red cross shows the unavailability of the on-board pilot ; overviews of the other pilot's screen are provided.

scenario, requiring a good understanding of the other pilot's ability to cope with a difficult situation. Five indicators (Heart Rate, emotions, Galvanic Response Skin, side stick input and gaze indicator) (see Fig. 9), are provided to help the ground assistant to adapt his communication according to the pilot's physiological state.

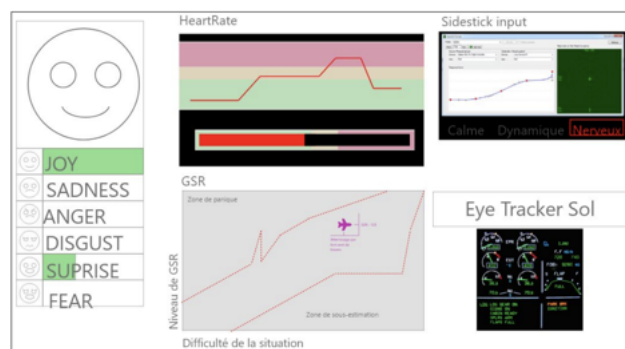


Figure 9. Prototype 3: Physiological state widgets. Left: an emotion indicator; Heart Rate and Galvanic Response Skin (GRS) indicators. Right: side stick input of the on-board pilot and gaze tracker that allow ground pilot to see what the on-board pilot is looking at.

The collaboration awareness explored in these prototypes (1, 2, 3), seem to us to be closely linked to awareness elements allowing a communication and coordination support. Indeed, thanks to these 2D on-screen awareness mechanisms (prototype 1, 2) and the sensor-mediated awareness mechanisms (prototype 3), the pilots may or may not engage in a conversation with the collaborator in a synchronous (voice, text chat) or asynchronous (mail) manner.

6.2 Exploring context awareness design

The aspect of context awareness explored in the following prototypes correspond to the view mode and allow pilots to be aware of the navigation of the collaborator in the shared map. The representation of the view mode has been explored in two

forms: a red view indicator and an overview indicator. The first design elements presented below are integrated into a shared map.

6.2.1 Prototype 4 and 5: view indicator

Inspired by the "radar views" of Gutwin et al. (1996), the red view indicator allows pilots to get the position of the view screen of the other pilot.

In the case of the synchronous shared map, both users can interact with the map and propose different rerouting solutions. To coordinate their action and communication, both pilots must have the same information displayed on the map.

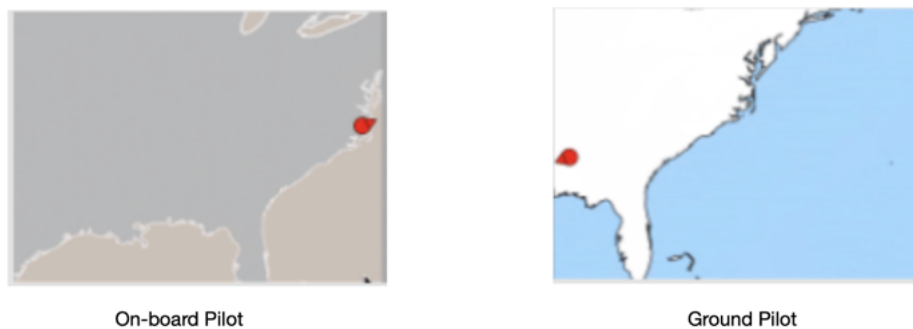


Figure 10. Prototype 4: First design of view indicator on remote synchronous shared map. Left: on-board pilot's screen with view indicator of ground pilot's screen in red. Right: Ground pilot's screen with view indicator of on-board pilot's screen.

To do this, we added a view indicator when the two pilots are not positioned at the same view on the map (see Fig. 10). A red dot with an arrow indicates the position of the other pilot's view (e.g. to the right slightly above). To get the same view as the other one the pilots have to click on the dot.

This tool is appropriate whenever the impact of context awareness can cause a disruption of social awareness. As this quote shows:

(P2.A): "and yes [...] so we both ended up with the QRH when one is supposed to fly [...]"

In this case, both pilots were forced to check the QRH (Quick Reference Handbook) because one of the pilots could not find the item concerning the "cabin smoke" situation (see section 5.4). Afterwards, they started a discussion about the supposed item but realized that it was not on the same page. We supposed that this loss of context awareness, highlighted by this quote, can be observed when consulting a map.

In the next prototype, an overview was added so that the pilot is able to see directly the view and the position of the other pilot's view (see Fig. 11).

The two awareness elements inform about the view of the other pilot and should help their coordination when talking about a precise area of the map.

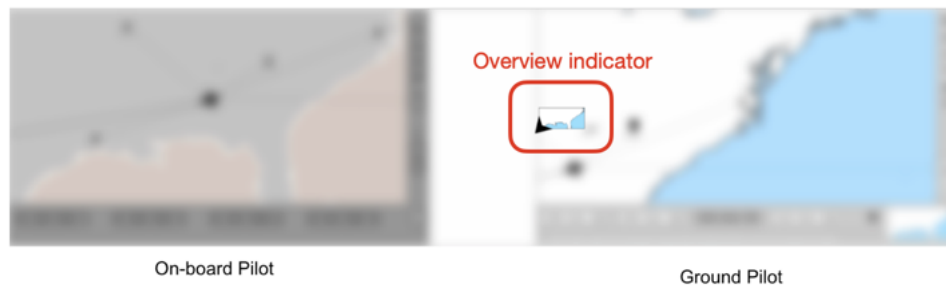


Figure 11. Prototype 5 : Second design of view indicator on remote synchronous shared map. On the left: on-board pilot' screen. On the right: Ground pilot' screen with view indicator of on-board pilot' screen.

6.3 Exploring workspace and social awareness design

The use of pointers for the map have been already used to support workspace awareness (Greenberg et al., 1996). It allows the pilot to know who, what, when, how he is manipulating the map. Indeed, we observed that in the interviews the accentuation of the verbalization allows to obtain a better workspace awareness. Although verbal communication will be maintained in the remote collaboration, we propose here a support to obtain information on who (name of the person on the mouse) is manipulating the map (what) when (screen synchronization) and how (using the mouse) (see Fig. 12).

6.3.1 Prototype 6: telepointers

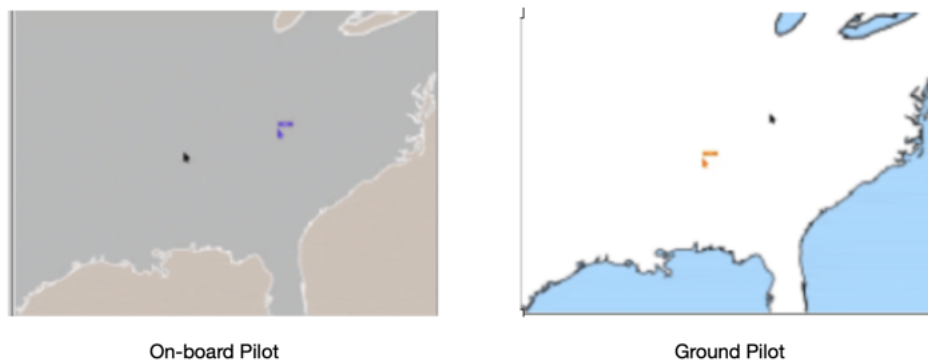


Figure 12. Prototype 6: Two pointers on shared maps (relaxed WYSIWIS) with on-board pilot (black pointer) and ground pilot (blue pointer) for on-board pilot screen. And on ground pilot' screen, an orange pointer that represent on-board pilot pointer and a black that represent his own pointer.

The idea of a telepointer seems to be a key element of the SiPO concept, as the two pilots explain:

(P3.B) : "uh [...] I know that at home we show the checklist a lot"
and

(P4.B) : " yeah you point to this checklist and you confirm [...]"

In addition, the color of the mouse indicates who has control over the shared map. When one of the pilots manipulates the map, the cursor changes color (here in blue, see Fig. 12) for the observer. When the other pilot is manipulating the map, there is a switch of roles within the collaboration tool and therefore of the cursor colors. Therefore, this design element improves social awareness by defining the role and activity in the tool.

We suppose that this element of workspace and social awareness can support coordination between the two pilots. Indeed, the two awareness elements, based on 2D on-screen awareness mechanisms (i.e. scroll bar, telepointers, chat tool or participant list (Hill and Gutwin, 2003)), inform about the actions provided in the shared map and allow pilots to coordinate their action on the shared map.

6.4 Exploring situation awareness

6.4.1 Prototype 7: Flight Timeline

The following prototype (see Fig. 13) provides an overview of the events occurring during the course of a flight in order to support situation awareness which might be necessary for a dedicated assistance. For instance, thanks to this timeline, a ground operator in a taxiing procedure would have a global awareness letting them better understand the behavior of the pilots, or assess the risk for them to miss a taxiway.

Possible examples of events include for instance switch to Automatic Pilot (AP), altitude changes, heading changes, ATC request, meteorological, systems failures or comments from the pilots. As the pilots explain :

(P3.B) "I thought to myself that it reminded me a little of the mechanic [...] the sailor who was at the back [...] I have this [...] with a global view even if he was overloaded because he had more work too"

and

(P4.B) : "no, but I was saturated [...] if he's suddenly reconnected like that, he needs the flight history, he needs to know where he is, the AP mode"

This timeline also gives access, by selecting a phase, to tools to monitor this phase more precisely (e.g. checklist, fuel, map, etc).

6.5 Feedback from pilots

The main aim of the exploration prototypes was to instantiate the concepts of awareness described in section 3. The designs were discussed with pilots to get feedback regarding awareness as well as usability. One of the most important feedback from the pilots was that the need for awareness is asymmetrical to the



Figure 13. Prototype 7: Flight Timeline displaying notable events during the different phases of a flight. Adding flight information is possible depending to the phase.

pilots' role (on-board or ground). Indeed, the addition of awareness elements for the pilot on board can be disruptive if the pilot does not choose to see these elements. This was confirmed by their request to decide themselves about the synchronicity of the map. In other words, pilots want to be able to choose between a relaxed WYSIWIS (What You See Is What I See) different views on their shared workspace (Greenberg, 1996) and a WYSIWIS (view shared between the users is identical)(Stefik et al., 1987). Concerning Prototype 3., pilots suggested that we should notify the pilot on the ground of abrupt changes and variations in these physiological values instead of continuous data. For the view indicator, the pilots seems to appreciate the idea with a preference for the Prototype 4. This preference is justified by its minimalism and the information transmitted. According to the pilots, prototype 5 is too large for the information given. Moreover, feedback was given on the color of the awareness elements (prototype 4, 5). For example, using red for the view indicator is inappropriate because this color is by convention a color associated with an urgent action to take. Finally, in Prototype 7, the pilots suggested that the timeline could be used as a reminder. This implies the addition of a possible shared editor inside the timeline. The shared editor could, according to the pilots, change the color of a phase, add icons, change the color of the icons or add markers.

7 Discussion and future work

In this section, we first discuss the contribution of this paper in refining the understanding of the awareness problem and secondly point out possible other problems in remote collaboration in SiPO context.

In our approach, we propose to refine the general awareness problem exposed by Lachter et al. (2014a), by applying awareness frameworks used for remote collaboration tools. These different proposals by Antunes et al. (2014), Schmidt (1998), and Gutwin et al. (1996) show us the range of awareness identifiable to support coordination, cooperation and communication in remote collaboration. In our case study and exploratory prototyping, we notice that the interpretation of awareness provided by Antunes et al. (2014) and Schmidt (1998) seems to be adapted to identify the need for awareness in future SiPO tools. Indeed, the

specification of social, workspace, context, situation and informal awareness in mutual awareness will allow us to specify the design of future remote collaboration tools in SiPO context.

Implicit interactions (Schmidt, 2000) have not been addressed in our case study nor in exploratory prototyping. A further step would therefore be to design awareness elements for this purpose. For example, to transmit implicit information of navigation interaction we can imagine a gradient of the indicator's red color to signify the type of interaction the pilot on board has on the shared map (unfinished, touch-and go, full interaction) (implicit interaction) (see Prototype 4).

To further refine the awareness problem, one of the points raised by the pilots is the relevance of awareness needs according to the role (on-board pilot and ground pilot). For instance, location awareness (to know where the ground pilot is geographically located) may be not necessary for the on-board pilot. To identify awareness according to roles, a possibility might be the utilization of the questionnaire by Antunes et al. (2014) (questionnaire to find out if the targeted awareness is implemented in the tool) and ask to the pilot if this awareness is necessary for the remote collaboration depending of their roles.

Another point to raise, is the need of awareness throughout the flight. We can suppose that the needs in terms of assistance will not be identical according to the phase of flight (variation of mental availability) or the situation (non-nominal, nominal).

The other aspects to be taken into account in the continuation of this study will finally be the impact of the automation and the reconfigurations to adapt to the various phases of flight, as well as the aspects of the collaboration between the operators on the ground dealing successively with the same flight.

8 Conclusion

In this paper, we proposed a specification of the awareness problem during remote collaboration in a new aeronautical context, the Single-Pilot Operation (SiPO). Thanks to our description of awareness frameworks applied in SiPO, our case study and our exploratory prototyping, we specified context awareness, workspace awareness, social awareness, informal awareness as key elements to design remote collaboration tools in this context. Finally, we illustrated them in the context of concrete scenarios and prototypes.

References

- Airbus (2022): 'Autonomous flight'. In: <https://www.airbus.com/en/innovation/autonomous-connected/autonomous-flight>.
- Antunes, P., V. Herskovic, S. F. Ochoa, and J. A. Pino (2014): 'Reviewing the quality of awareness support in collaborative applications'. *Journal of Systems and Software*, vol. 89, pp. 146–169.

- Beaudouin-Lafon, M. and A. Karsenty (1992): 'Transparency and awareness in a real-time groupware system'. In: *Proceedings of the 5th annual ACM symposium on User interface software and technology*. pp. 171–180.
- Becquet, V., C. Letondal, J.-L. Vinot, and S. Pauchet (2019): 'How do Gestures Matter for Mutual Awareness in Cockpits?: Disclosing Interactions through Graphical Representations'. In: *Proceedings of the 2019 on Designing Interactive Systems Conference*. p. 593–605, ACM.
- Bellotti, V. and A. Sellen (1993): 'Design for privacy in ubiquitous computing environments'. In: *Proceedings of the Third European Conference on Computer-Supported Cooperative Work 13–17 September 1993, Milan, Italy ECSCW'93*. pp. 77–92.
- Bilimoria, K. D., W. W. Johnson, and P. C. Schutte (2014): 'Conceptual framework for single pilot operations'. In: *Proceedings of the international conference on human-computer interaction in aerospace*. pp. 1–8.
- Billings, C. E. (1997): 'Aviation Automation the Search for a Human Centered Approach Lawrence Erlbaum Associates'. Inc., Mahwah, NJ.
- Bohn, R. E. (2010): 'How flying got smarter'. In: *2010 IEEE Intelligent Vehicles Symposium*. p. 682–687, IEEE.
- Brandt, S. L., J. Lachter, V. Battiste, and W. Johnson (2015): 'Pilot Situation Awareness and its Implications for Single Pilot Operations: Analysis of a Human-in-the-Loop Study'. *Procedia Manufacturing*, vol. 3, pp. 3017–3024.
- Bravo, C., R. Duque, and J. Gallardo (2013): 'A groupware system to support collaborative programming: Design and experiences'. *Journal of Systems and Software*, vol. 86, no. 7, pp. 1759–1771.
- Carroll, J. M., D. C. Neale, P. L. Isenhour, M. B. Rosson, and D. S. McCrickard (2003): 'Notification and awareness: synchronizing task-oriented collaborative activity'. *International Journal of Human-Computer Studies*, vol. 58, no. 5, pp. 605–632.
- CleanAviation (2022): 'At the sharp end: Innovative cockpit tech points the way'. In: <https://clean-aviation.eu/media/results-stories/at-the-sharp-end-innovative-cockpit-tech-points-the-way>.
- Council, N. R. et al. (2014): *Autonomy research for civil aviation: toward a new era of flight*. National Academies Press.
- Dao, A.-Q. V., K. Koltai, S. D. Cals, S. L. Brandt, J. Lachter, M. Matessa, D. E. Smith, V. Battiste, and W. W. Johnson (2015): 'Evaluation of a recommender system for single pilot operations'. *Procedia Manufacturing*, vol. 3, pp. 3070–3077.
- Duchevet, A., G. Granger, J.-P. Imbert, T. De La Hogue, A. L. R. Vasquez, J. N. Sanchez, J. M. M. Garcia, S. Bonelli, and A. Ferreira (2020): 'Toward a Non Stabilized Approach assistant based on human expertise'. In: *ICCAS 2020, 1st International Conference on Cognitive Aircraft Systems*.
- Endsley, M. R. (2017): 'Toward a theory of situation awareness in dynamic systems'. In: *Situational awareness*. Routledge, pp. 9–42.
- European Union Aviation Safety Agency, E. (2021): 'Horizon Europe Project: Extended Minimum Crew Operations – Single Pilot Operations – Safety risk assessment framework'.
- Greenberg, S. (1996): 'A fisheye text editor for relaxed-WYSIWIS groupware'. In: *Conference Companion on Human Factors in Computing Systems*. pp. 212–213.

- Greenberg, S. and C. Gutwin (2016): 'Implications of we-awareness to the design of distributed groupware tools'. *Computer Supported Cooperative Work (CSCW)*, vol. 25, no. 4, pp. 279–293.
- Greenberg, S., C. Gutwin, and M. Roseman (1996): 'Semantic telepointers for groupware'. In: *Proceedings sixth Australian conference on computer-human interaction*. pp. 54–61.
- Gross, T. (2013): 'Supporting effortless coordination: 25 years of awareness research'. *Computer Supported Cooperative Work (CSCW)*, vol. 22, no. 4-6, pp. 425–474.
- Gutwin, C., S. Greenberg, and M. Roseman (1996): 'Workspace awareness in real-time distributed groupware: Framework, widgets, and evaluation'. In: *People and computers XI*. Springer, pp. 281–298.
- Hill, J. and C. Gutwin (2003): 'Awareness support in a groupware widget toolkit'. In: *Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work*. pp. 258–267.
- Huddleston, J., D. Harris, D. Richards, S. Scott, and R. Sears (2015): 'Dual pilot and single pilot operations—hierarchical task decomposition analysis of doing more with less'. In: *International Conference on Engineering Psychology and Cognitive Ergonomics*. pp. 365–376.
- Johansen, R., D. Sibbet, S. Benson, A. Martin, R. Mittman, and P. Saffo (1991): *Leading business teams: How teams can use technology and group process tools to enhance performance*. Addison-Wesley Longman Publishing Co., Inc.
- Koch, M. and J. Koch (2000): 'Application of frameworks in groupware—The iris group editor environment'. *ACM Computing Surveys (CSUR)*, vol. 32, no. 1es, pp. 28–es.
- Lachter, J., V. Battiste, M. Matessa, Q. V. Dao, R. Koteskey, and W. W. Johnson (2014a): 'Toward single pilot operations: the impact of the loss of non-verbal communication on the flight deck'. In: *Proceedings of the International Conference on Human-Computer Interaction in Aerospace - HCI-Aero '14*. p. 1–8, ACM Press.
- Lachter, J., S. L. Brandt, V. Battiste, S. V. Ligda, M. Matessa, and W. W. Johnson (2014b): 'Toward single pilot operations: developing a ground station'. In: *Proceedings of the International Conference on Human-Computer Interaction in Aerospace - HCI-Aero '14*. p. 1–8, ACM Press.
- Lachter, J., S. L. Brandt, V. Battiste, M. Matessa, and W. W. Johnson (2017): 'Enhanced ground support: lessons from work on reduced crew operations'. *Cognition, Technology Work*, vol. 19, no. 2–3, pp. 279–288.
- Letondal, C., J.-L. Vinot, S. Pauchet, C. Boussiron, S. Rey, V. Becquet, and C. Lavenir (2018): 'Being in the sky: Framing tangible and embodied interaction for future airliner cockpits'. In: *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*. pp. 656–666.
- Ligda, S. V., U. Fischer, K. Mosier, M. Matessa, V. Battiste, and W. W. Johnson (2015): 'Effectiveness of advanced collaboration tools on crew communication in reduced crew operations'. In: *International conference on engineering psychology and cognitive ergonomics*. pp. 416–427.
- Mackay, W. E. (1999): 'Media spaces: Environments for informal multimedia interaction'. *Computer Supported Co-operative Work*, vol. 7, pp. 55–82.
- Niermann, C. and D. Kügler (2021): 'Next Generation Intelligent Cockpit—Single-pilot operations' potential to increase flight safety'. *ECAC News*, no. 74, pp. 25–28.

- Paz Goncalves Martins, A., J. Lieb, M. Friedrich, S. Bonelli, M. Celori, and A. de Bortoli (2021): 'Toward single pilot operations: A conceptual framework to manage in-flight incapacitation'. *11th SESAR Innovation Days: Inspiring Long-Term Research in the Field of Air Traffic Management, SIDs 2021*.
- Röcker, C. (2012): 'Informal communication and awareness in virtual teams'. *Communications in Information Science and Management Engineering*, vol. 2, no. 5, pp. 1–15.
- Rodden, T. (1996): 'Populating the application: A model of awareness for cooperative applications'. In: *Proceedings of the 1996 ACM conference on Computer supported cooperative work*. pp. 87–96.
- Schmid, D. and N. A. Stanton (2019): 'A future airliner's reduced-crew: modelling pilot incapacitation and homicide-suicide with systems theory'. *Human-Intelligent Systems Integration*, vol. 1, no. 1, pp. 27–42.
- Schmid, D. and N. A. Stanton (2020): 'Progressing Toward Airliners' Reduced-Crew Operations: A Systematic Literature Review', pp. 25.
- Schmidt, A. (2000): 'Implicit human computer interaction through context'. *Personal technologies*, vol. 4, no. 2, pp. 191–199.
- Schmidt, K. (1998): 'Some notes on mutual awareness'. *COTCOS, CTI, DTU, Copenhagen, Denmark*.
- Schmidt, K. (2002): 'The problem with awareness': Introductory remarks on awareness in CSCW". *Computer Supported Cooperative Work (CSCW)*, vol. 11, no. 3-4, pp. 285–298.
- Snowdon, D., E. F. Churchill, and A. J. Munro (2000): 'Collaborative virtual environments: digital spaces and places for CSCW'. *Collaborative Virtual Environments (2000)*, pp. 1–34.
- Stefik, M., D. G. Bobrow, G. Foster, S. Lanning, and D. Tatar (1987): 'WYSIWIS revised: Early experiences with multiuser interfaces'. *ACM Transactions on Information Systems (TOIS)*, vol. 5, no. 2, pp. 147–167.

Permissions

If the paper is accepted for publication, a separate transfer-of-copyright form must be signed and submitted as a PDF or TIFF file, by ordinary mail, or by fax no later than the deadline for submission of camera-ready manuscripts. It is the obligation of the author(s) to obtain written permission for quotations from unpublished material, for all quotations in excess of 250 words in one extract or 500 words in total from any work still in copyright, and for the reprinting of illustrations or tables from unpublished or copyrighted material.

Woosuk Seo¹, Joonyoung Jun², Minki Chun³, Hyeonhak Jeong³, Sungmin Na², Woohyun Cho¹, Saeri Kim⁴, Hyunggu Jung³ (2022): *Toward an AI-assisted Assessment Tool to Support Online Art Therapy Practices: A Pilot Study*. In: *Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies -Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_ep04*

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_ep04

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

Toward an AI-assisted Assessment Tool to Support Online Art Therapy Practices: A Pilot Study

Woosuk Seo¹, Joonyoung Jun², Minki Chun³, Hyeonhak Jeong³,
Sungmin Na², Woohyun Cho¹, Saeri Kim⁴, Hyunggu Jung³

¹University of Michigan, ²University of Washington, ³University of Seoul,

⁴University of the Arts London

Corresponding Author: Hyunggu Jung (hjung@uos.ac.kr)

Abstract. Artificial intelligence (AI) has been widely used to assist art therapists with artwork assessments by providing objective information. While prior studies showed that AI-assisted tools are feasible to improve drawing analysis in in-person art therapy practices, the use of those tools in online art therapy is still under-examined. To fill the gap, we created a prototype of an AI-assisted tool for online therapy in a House-Tree-Person (HTP) test scenario and ran lab-based usability sessions with 10 art therapists in which they used our proposed prototype to complete predefined tasks. We then conducted semi-structured interviews with the participants to understand their acceptance and concerns about the prototype. The findings revealed the unique needs of art therapists and opportunities of using AI-assisted tools to improve online art therapy practices. Based on these findings, we suggest implications for creating AI-assisted tools that meet specific needs of art therapists in online therapy sessions, and further discuss future directions of research about AI-assisted tools for art therapists in online settings.

1 Introduction

During the pandemic of COVID-19, art therapy has supported people in overcoming emotional, psychological, and social issues. Through active art-making such as drawing, sculpturing, or painting, people foster self-esteem, develop emotional resilience, improve social skills, reduce conflict, and resolve distress. To offer art therapy in the pandemic while keeping social distance for safety, the majority of art therapists have to facilitate online therapy sessions (Zubala and Hackett, 2020). However, art therapists face many challenges while conducting online art therapy sessions (Zubala and Hackett, 2020; Snyder, 2021; Power et al., 2021; Choudhry and Keane, 2020). These challenges occur in viewing clients' art-making processes, incorporating art materials, and finding guidelines for the online sessions (Zubala and Hackett, 2020; Choudhry and Keane, 2020). Above all, many art therapists have faced various challenges related to using technology. Those challenges include a steep learning curve to deliver online therapy practice, a lack of knowledge or skills to manage different equipment, and issues with broadband connectivity. Especially, due to the sudden shift to online therapy practices, art therapists might have a shortage of time to research or learn about using digital technology as a tool to connect with clients remotely. While the challenges and concerns are related to the sudden shift to online settings, technology still plays an important role in art therapists' experience with online sessions.

Even before online art therapy became prevailing due to the pandemic, researchers demonstrated benefits of using innovative technology for art therapy from personal computers (Parker-Bell, 1999) to virtual reality (Hacmun et al., 2018). First, communication technologies enhance interactions between clients and art therapists in remote settings. For instance, recent studies show art therapists facilitate online art therapy through applications, such as Zoom¹, WhatsApp², or Skype³ (Collie et al., 2017; Levy et al., 2018; Lloyd and Usiskin, 2020; Zubala and Hackett, 2020). Furthermore, technologies can be used to support art therapists with analyzing clients' artwork. Particularly, artificial intelligence (AI) is used for analyzing artwork in art therapy, such as identifying the main color in a drawing (Kim, 2008). Nevertheless, existing AI-assisted tools for art therapists are mostly designed to support in-person sessions or focus on the clients' experience in online therapy sessions. Unlike in an in-person environment, art therapy in online settings may impose various challenges when art therapists use AI-assisted tools along with other devices for conducting online art therapy. It is important to acknowledge such specific challenges of art therapists in online art therapy.

Extending the line of research on the development of technology for improving art therapy practices, we aimed to fill the gap in the literature about AI systems in healthcare contexts. Thus, we developed and evaluated AI-assisted assessment

¹ <https://www.zoom.us/>

² <https://www.whatsapp.com/>

³ <https://www.skype.com/>

tools with 10 art therapists. We identified their acceptance and concerns about the prototype and revealed potential benefits of using AI-assisted tools to support online art therapy practices. We then suggest implications for creating an AI-assisted tool that meet specific needs of art therapists in online therapy sessions. Hence, our pilot study makes the following contributions to the CSCW community: (1) we provide a preliminary empirical understanding of how art therapists perceive an AI-assisted assessment tool for online art therapy practices; (2) we identify needs of art therapists when using an AI-assisted assessment tool for online therapy practices; and (3) we present implications for AI-assisted tools to improve the quality of the online drawing-based assessments.

2 House-Tree-Person Test

In our study, we chose House-Tree-Person (HTP) test for prototype development since it is widely used in therapy sessions or clinic consultations (Rohail, 2015; Yu et al., 2016; Groth-Marnat and Roberts, 1998). The HTP test is a projective drawing technique developed by a psychologist named John Buck (1948). While HTP test was originally developed to identify personal traits and intelligence, it has been used as a tool for an initial assessment in psychotherapy or mental health services. For instance, HTP test was used to assess risk factors of personality of patients with mental disorders (Rohail, 2015). In the HTP test, a client receives a short and abstract instruction to draw a house, a tree, and the figure of a person (e.g., “Draw any house”). During and after the client draws, the administrator of the test (e.g., therapist, psychiatrist) asks the client to describe the drawing (e.g., “Who lives in this house?”). Through such interactions, the administrator is allowed to investigate symbolic meanings of drawings that project the client’s personal traits. For instances, the size and number of door and windows of a house drawing represent the client’s relationship to the outside world and socialization. Although the interpretation of drawings may be affected by the administrator’s individual knowledge and experiences, Buck’s manual provides a general guide for facilitating a HTP test and identifying meaningful information from drawings (Buck, 1948). Thus, we adopted Buck’s HTP test manual and developed a prototype that provides art therapists with information about drawings in online therapy sessions.

3 Related Work

3.1 Traditional Technology-assisted Art Therapy

Prior studies showed how technology supports art therapists in conducting both in-person and online art therapy sessions. Researchers in medical and HCI communities have developed and evaluated technologies that support various aspects of art therapy, including new art media and online group therapy sessions.

Some prior studies focused on technology-assisted art therapy practices that involves digital media. Instead of asking clients to draw objects on paper with pencils, technology allows art therapists to utilize digital collage making (Diggs et al., 2015), animation (Austin, 2009), or digital photography (Atkins, 2007). Those prior works showed that the diversity of digital media may enable art therapists to assess clients' emotional and psychological distress with various approaches. Furthermore, some researchers have been seeking the opportunities for technology to support online art therapy (Hacmun et al., 2018; Datlen and Pandolfi, 2020; Collie et al., 1998; Collie and Cubranić, 2002; Jones et al., 2014). These prior studies mostly focused on developing tools to improve communication between a therapist and clients or among clients in online group therapy sessions by providing a virtual space to share artwork (Datlen and Pandolfi, 2020) or enabling collaborative drawing (Jones et al., 2014).

3.2 Artificial Intelligence (AI) for Assessment Practices and Art Therapy

Artificial Intelligence (AI) has been supported assessment practices in various contexts, including healthcare and online learning. In a healthcare domain, researchers developed AI systems that predict clients' prognoses or recommend personalized treatment (Cohen et al., 2020; Yao et al., 2020; Schwartz et al., 2021; Gómez Penedo et al., 2021). While Schwartz et al. (2021) developed a treatment selection model to recommend clients' optimal treatment, Goldberg et al. (2021) evaluated the feasibility of using AI-assisted tools to assess therapists' interpersonal skills that are fundamental to psychotherapy. Similarly, in an online learning domain, prior studies focused on investigating how AI systems support the assessment of learners' understanding and competency (Coghlan et al., 2021; Hamdy et al., 2021). While Coghlan et al. (2021) discussed ethics of AI-assisted online proctoring technologies, Hamdy et al. (2021) evaluated the feasibility, acceptability, reliability and validity of an online clinical encounter tool designed to assess aspects of medical students' non-psychometric competencies. Considering as psychotherapy, art therapy has also been influenced by the advancement of artificial intelligence (AI). Particularly, AI has supported art therapists in interacting with their clients or analyzing clients' artwork. While AlSadoun et al. (2020) proposed an AI-assisted art therapy system that uses technology to enhance communication between therapists and clients, Kim and colleagues designed AI systems to evaluate the color-related information from clients' artwork, including the number of colors used and the length of edges between colors (Kim et al., 2007; Kim, 2008; Kim and Hameed, 2009; Kim et al., 2009, 2008). Moreover, in a recent study, Kim et al. (2021) proposed a novel deep learning model that may increases efficiency and objectivity of art therapy practices by identifying symbolic features from HTP drawings.

3.3 Limitations of Prior Studies

In brief, prior studies showed the opportunities for AI to improve the quality of art therapy practices, such as analyzing color elements of drawings (Kim et al., 2007; Kim, 2008; Kim and Hameed, 2009) or providing information about symbolic features of drawings (Kim et al., 2021). Nevertheless, those studies mostly investigated how to improve AI's efficiency and effectiveness to support art therapists in decision-making. The perspective and experience of art therapists with the AI-assisted tool are still understudied. In particular, little is known about how existing AI-assisted tools would be integrated into online art therapy practices that may impose unique challenges to art therapists. Hence, our study extends the understanding of AI-assisted tools' integration with online art therapy practices and provides empirical findings on how an AI-assisted assessment tool supports art therapy practices.

4 Method

We developed an AI-assisted prototype and conducted lab-based usability sessions with 10 art therapists to evaluate the prototype. In each session, each participant was asked to learn about the prototype, performed the predefined tasks, and participated in an interview. By following the procedure using the Wizard of Oz approach (Dahlbäck et al., 1993). Each participant interacted with our proposed prototype that offer manually typed information about each element (e.g., window) of drawings, rather than the information automatically detected and provided by AI models. Also, we did not ask about or measure participants' knowledge about HTP test since it was not the aim of study.

4.1 Prototype Development

Based on the findings from prior studies about challenges and needs of art therapists, we were motivated to develop a prototype that supports the assessment process in online art therapy sessions. Particularly, we focused on online HTP test as it has been widely used in in-person sessions. The development of our prototype was done in 5 steps: (1) creating an HTP test checklist following John Buck's manual, (2) collecting sample House, Tree, and Person drawings from actual tests conducted by John Buck (1950), (3) creating information about the drawings that the user would seek for assessment, based on the Buck's manual (Buck, 1948), (4) creating a server to run our prototype on with Flask web framework and Python, and (5) creating the prototype with HTML and Javascript. Our prototype supports art therapists navigating information about the selected drawing and analyzing the drawing (see Figure 1). Specific information about an element (e.g., window) or overview information about each drawing are described in to three aspects: detail, perspective, and proportion (Buck, 1948). Particularly, art therapists are allowed to click red bounding boxes around elements that are detected by AI (see Figure 2).

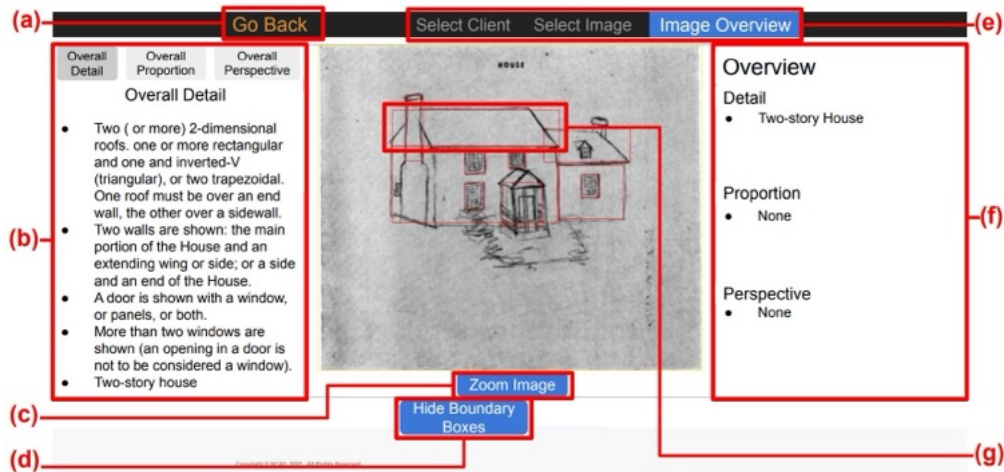


Figure 1. The prototype with main features are highlighted: (a) Back button to previous page; (b) Overall information of the image; (c) A button for displaying the enlarged image in a pop-up window; (d) A button for viewing the original image where specific items are not marked with red bounding boxes; (e) A progress indicator of the prototype; (f) The detail, proportion, and perspective of the selected element; and (g) Clickable bounding box that shows information about specific element of the drawing.

Once clicked, the prototype provides art therapists with information about the specific element.

4.2 Data Collection

We recruited 10 registered art therapists (see Table 1). The eligibility criteria for the participants are as follows: (1) must be over 18, (2) must be able to understand English, (3) must be a registered art therapist with titles (e.g., ATR, ATR-P, ATR-BC, and LCAT), and (4) must have conducted at least one online art therapy session in the past year. To identify eligible study participants, we obtained a list of registered art therapist on the websites of institutions credentialed by the American

P#	Gender	Age	Race/ethnicity	Year(s) since certificated	Learned HTP
P1	M	39	White	2.5	Yes
P2	F	46	White	16.0	No
P3	F	61	White	5.5	Yes
P4	N/A*	53	N/A*	22.5	Yes
P5	F	27	Hispanic	1.0	Yes
P6	F	57	White	26.0	No
P7	F	42	Asian, Hawaiian, & Hispanic	8.5	Yes
P8	F	42	White	10.5	Yes
P9	F	68	White & Hispanic	44.0	Yes
P10	F	38	White	8.0	Yes

Table I. The demographics information of the interview participants. * P4 preferred not to describe gender and race/ethnicity. Two participants (P2, P6) did not learn the assessment process of HTP tests, but it did not affect their completion of tasks in our usability sessions since the tasks did not require prior knowledge about HTP test.

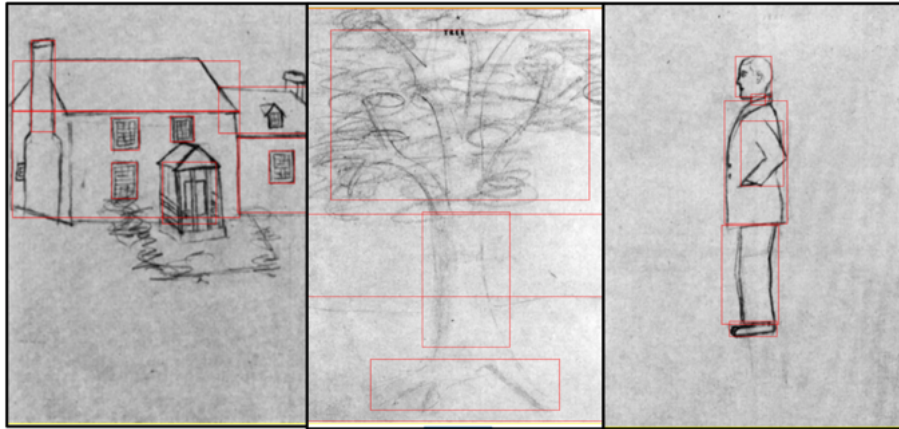


Figure 2. The three drawings: House (left), Tree (middle), and Person (right). For the Tree drawing, it was challenging for some participants to differentiate the “Branch” and “Foliage” as their bounding boxes seem to be overlapped to each other. Red bounding boxes indicate the element that AI detects. When each bounding box is clicked, AI provides art therapists information about the specific element.

Art Therapy Association. We also searched art therapists on Facebook, YouTube, LinkedIn, and Reddit by using the search queries (e.g., “art therapy,” or “art therapist”). The participant age ranges from 27 to 68 ($M = 47.3$, $SD = 12.3$). The years since they get certificated in art therapy ranges from 1 to 44 ($M = 14.4$, $SD = 13.2$). All 10 participants have worked with adolescents (age 14-19) and adults (age 20-64) clients, while some of them specifically worked with preschool children (age 2-5), school-aged children (age 6-13), and/or seniors (age 65+). Eight of the participants have learned HTP test. This study was approved by our university’s Institutional Review Board prior to data collection.

To collect data, we conducted a usability session with each participant as follows: (1) We introduced our prototype to participants, including how main features of our prototype work, and how AI analyze drawings; (2) We asked participants to try out our prototype by completing demo tasks so that they can understand how the prototype works and what information it provides; (3) We asked participants to use our prototype for performing predefined tasks; and (4) We conducted a semi-structured interview with each participant. The predefined tasks were: (1) searching specific information of each drawing and type it to a Google Doc, (2) searching overall information of each drawing and type it to a Google Doc. For example, the participants were asked to navigate our prototype to find detail information of roof from House drawing and overall perspective information from Tree drawing. These tasks were inspired by Buck’s manual (Buck, 1948) that guide therapists to seek information from each element (e.g., door) from each drawing and to take notes for further assessment if necessary.

Next, we conducted post-test interviews that lasted 25-30 minutes. Participants gave verbal consent to interview recording and the use of anonymous data at the beginning of the interviews. They acknowledged that their participation was

voluntary. Interview questions pertained to their experience regarding using the prototype and their opinions about potential features and use of the prototype. The sample questions include “*What are your overall impressions of the prototype?*”, “*How do you feel about the information provided by AI*”, and “*How would you integrate our prototype into your online therapy practice?*”. To ensure the anonymity of all participants, we removed all personally identifiable information and assigned each participant a unique identifier (e.g., P1, P2). Each participant was compensated with a \$20 gift card for their participation in this study. All interviews were recorded and transcribed automatically by Zoom, and the transcripts were reviewed by research team members to fix any errors.

4.3 Data Analysis

We analyzed the interview data using grounded theory (Strauss and Corbin, 1997). Six out of the authors separately coded the transcripts. The entire research team then conducted an affinity diagram (Holtzblatt and Beyer, 1997) on Miro⁴ to identify key themes, insights, and patterns that appeared repeatedly in the collected qualitative data. To resolve discrepancies between researchers, the team had weekly meetings. , we identified how art therapists perceive an AI-assisted tool and what they need to integrate the system into their online therapy practices. The research team categorized and refined emerging themes (e.g., perceived benefits of using our prototype, challenges in adopting the prototype into current practices, and potential features to be added to the prototype).

5 Results

By analyzing interview data, we found four salient themes: benefit and challenge of using the bounding boxes, evaluation on prototype features, evaluation on information provided by the prototype, and potential benefits as an educational tool for art therapy students. Despite of diversity of our participants, these themes emerged across the whole range of work experience and age.

5.1 Roles of the Bounding Boxes

An important user interface element of our prototype is the bounding boxes that visualize selected elements on drawings. Our participants reported a major benefit and a significant challenge in utilizing the bounding boxes. For the benefit, four participants reported that the bounding boxes allowed them to analyze subtle information about elements (e.g., proportion of arm of Person) as such information is hard to eye-measure. For example, P3 emphasized that bounding boxes were helpful to specific elements of drawings: “*I really like the ability to highlight boundaries, because I definitely have certain things that I look for in the House*

⁴ <http://www.miro.com/>

[drawing]. I was able to highlight the windows and I really think that the windows had a lot of information in them. [...] I can really look at certain features more closely.” (P3). On the other hand, three participants were confused with the prototype’s features, particularly with the bounding boxes on the drawings. They pointed out that the bounding boxes misled them because of the lack of sufficient explanation about the bounding boxes and AI detection. For instance, P8 assumed the red bounding boxes on the sample drawings were drawn by the client: *“I assumed that they were like yellow or red boxes, or something the client drew. [...] I think it would be necessary to explain that that box is being put on by the system so that I don’t interpret that the client suddenly put a red square.”* (P8). Particularly, two participants expressed both satisfaction and concerns for the bounding boxes. Their ambivalent perceptions were identified when they were asked about their concerns about the prototype and potential features to be added to the prototype.

5.2 Potential Prototype Features

In addition to the bounding boxes, our prototype provided the participants with several features, including displaying different categories (e.g., proportion) of information about the drawing elements. However, they expressed their desire for additional features for the prototype. For example, eight participants were interested in using the prototype for other assessment techniques, such as Person in the Rain, and Person Picking an Apple. Participants wanted to use the prototype for the assessments that they felt more comfortable with. The suggested assessments include Person in the Rain, and Person Picking an Apple. One of the common characteristics of these assessment techniques, including the HTP test, is that a manual is necessary for therapists to conduct a assessment with their clients. Moreover, four participants stated that they were concerned about the prototype’s lack of the scalability to adjust to therapists’ individual practices. Two out of the four participants described they were eager to follow their own manuals because our proposed prototype only addressed Buck’s HTP manual considered as an original manual. Even if they use structured assessment techniques (e.g., the HTP test), our participants tended to have their own technique of facilitating assessments: *“I think the House Tree Person is a lot different for art therapists. Everyone kind of has their own way of doing it, even though there’s like a very structured way to conduct it.”* (P5). This quote indicates therapists’ desire to have the ability to customize AI systems to be integrated with their individual therapy practices, even though there are structured manuals with fixed classification values and information about drawings.

5.3 Essential Clients’ Data for Assessing Drawings

We found that therapists in our study wanted to obtain additional information about clients for our AI-assisted prototype. Although our prototype displayed the

client's background information such as demographic information, four participants expressed their concerns that the prototype did not incorporate client's emotions or behaviors into AI-provided information during the session. For example, P4 stated that the prototype did not consider the diverse background of the clients and difficult to observe how a client created his/her artwork. *"It seems like everybody is going to meet things that are going to mean the same thing for every person and that's not true. It's not taking into account all of the intersections of a person's diversity, it doesn't take into account cultural all kinds of different background information that is very unique to each individual and how they all intersect and you know. One thing like a hole in a tree doesn't mean the same thing for every person."* (P4). Lack of additional information about the clients made study participants feel the full assessment conducted by AI would not be truthful as the AI-assisted tool did not consider interactions between the clients and their drawings. This limitation may also cause a potential consequence leading to misinterpretation of clients' drawings. For instance, the prototype may not correctly detect what the clients intended to draw. Then, art therapists may misinterpret the drawings. Acknowledging the limitation, P5 suggested the prototype should consider clients' intentions: *"If a client draws like a star on a house or something and the AI doesn't pick that up as a window let's say the client decides it's a window and maybe there could be an option, where we[art therapists] could manually kind of put it in and say like this is a window."* (P5).

5.4 Potential Benefits as an Educational Tool for Art Therapy Students

In addition to improving the efficiency of online art therapy, we noticed that our prototype has a potential to be used as an educational tool to support art therapy students. For instance, P3 mentioned the opportunity of an AI-assisted tool for art therapy students as a standardized tool to supplement their subjective interpretation of drawings: *"What I want the students to understand is that they might be in a setting where they don't always get to do our therapy as much as they would like to [...] They could even present this to you know their colleagues their supervisors and talk about how this is a standardized tool and it's interpreted with the help of AI [...] I think this could be seen as a really you know standardized valid tool."*(P3). However, P9 was worried about letting students use the prototype because of their potential dependency on AI-assisted tools: *"So, my concern would be that a student would use your prototype and not really look at that information and think about that information themselves [...] They would just go the easy path, instead of really trying to look at the image themselves and come up with those conclusions themselves."* (P9). P9's concern indicates potential negative influences of the prototype as an educational tool on students' assessment skills, instead of helping them develop their skills.

6 Discussion

In this section, we discuss how to better design an AI system for art therapists in online therapy sessions. The findings of this interview study revealed that our AI-assisted prototype needs to be improved in three aspects. Each aspect indicates the needs of our participants, including understanding AI's functions, extending the application of the prototype beyond online therapy sessions, and considering clients' input for an AI-assisted assessment tool. We also discuss directions for designing potential collaborative AI systems for online art therapy sessions.

6.1 Visual Cues for Explaining AI's Features

A major concern that our participants had was that the prototype did not sufficiently explain about AI functions. Although the participants mentioned that our prototype was useful to search specific detailed information on drawings, they also expressed their desire to have more explanations about what the AI-assisted tool can do. Particularly, four participants were not satisfied with the bounding boxes on drawings due to the limited explanation about how our prototype interprets the drawings. Aligning with prior studies about explainable AI (e.g., Chromik (2020)), our finding indicates the need for sufficient explanation about how AI works, where it derives information from, and what it shows to therapists. To meet such need for sufficient explanation about medical AI systems, prior studies have presented design implications. Especially, our findings resonate with the results from a study by Xie et al. (2020) that reported the development and feasibility of an AI system for physicians. By conducting co-design activities with physicians to formulate an AI system (CheXplain), that provides physicians with analysis information of chest x-ray images, the authors discovered the unique needs of physicians when physicians use the system to analyze images to make medical decisions. The authors also presented design recommendations for developing medical AI systems that can be understandable to physicians. A key recommendation was allowing physicians to control the range of AI-provided information that they would want it to be explained, from just showing the results to revealing in-depth explanations about the results. This recommendation is also applicable to the online art therapy context but with more emphasis on the visual cues such as colors and shapes of bounding boxes that helps art therapists to control AI-provided information that may need explanation. As shown in our findings, some therapists were confused with the bounding boxes on drawings. Art therapists' confusion related to bounding boxes on drawings may reduce if there are appropriate visual explanations about the AI-provided information. For example, an AI system may allow art therapists to control the number of elements with bounding boxes shown on the drawings or change the color of bounding boxes that clients have not used.

6.2 Potential Opportunities to Support Art Therapy Education Beyond Online Therapy Sessions

While our prototype was initially designed to assist online therapy practices, our findings revealed opportunities for an AI-assisted assessment tool to support therapists beyond online therapy sessions. Our participants emphasized potential use of an AI-assisted assessment tool for developing art therapy students' assessment skills and knowledge. Similar to how AI systems are used for educating medical students (e.g., chatbot for practicing communication skills of medical students (Shorey et al., 2019)), AI systems may be integrated with art therapy curriculum to improve learning achievements. In particular, art therapy students may benefit from AI systems to develop their skills to conduct online drawing-based assessments as they would get supplementary subjective information about drawings from AI systems. Also, AI systems may provide students with personalized learning environment. For instance, an AI system can be used to collect data about a student's behaviour during online sessions. The system may use the collected data to predict the student's behaviour and provide relevant analysis information from the manual to help the student develop analysis skills. However, it might be crucial to consider a potential issue with using AI systems for developing students' assessment skills. As P9 mentioned, students may have over-dependency on AI systems such as excessively relying on the results from AI systems, rather than developing their own skills to assess elements from drawings. Although dependency on AI systems may improve the efficiency of teaching environment as information about drawings would be automatically provided to students, over-dependency can be a major issue for an AI-assisted assessment tool with educational purpose that focus on developing students' skills. To reduce such over-dependency, we suggest that developers adopt the cognitive forcing interventions (e.g., AI offers information only when a button is clicked (Buçinca et al., 2021)). Such interventions may allow students to process information from the drawings before seeing the results from an AI-assisted educational system. In sum, even though AI-assisted assessment tools can potentially support art therapy curriculum, they may also hinder developing students' assessment skills. Further studies are necessary to explore more potential opportunities of AI-assisted assessment tools to be used for educational purpose.

6.3 Client-provided Information to Facilitate Therapist-Client Interactions

To better support art therapists' online interactions with their clients, we suggest that researchers add features that enable AI-assisted tools to obtain client-provided information. The client-provided information include data about clients who drew the drawings, their intentions for drawing specific figures, and their representations of certain elements (e.g., a star-shaped window [P5]). AI-assisted tools may use such client-provided information that allows art therapists to analyze the clients'

drawings. For instance, if the AI-assisted tool detects that the client did not draw a chimney, it notifies art therapists to ask the client further questions about it. In that case, an AI-assisted tool detects and highlights the elements that do not match with its manual (e.g., Buck's HTP manual). It allows art therapists and clients to talk about those specific elements. The visual cues (e.g., bounding boxes) on the drawings would make clients more engaged with the conversations since they know what art therapists would like to talk about. While considering clients as a user, it is also essential to address potential concerns of clients that could be similar to other patients in remote patient-provider interactions. For example, recent CSCW studies showed that some patients were worried about the lack of empathy and care quality (Bhat et al., 2021), and about providers' dismissal of symptoms during online medical appointments (Raj et al., 2022). Furthermore, prior studies showed the importance of integrating patients' preferences to AI-assisted tools for healthcare professionals. For instance, Jacobs et al. (2021) found the significance of including patient preferences in AI decision supporting tools for healthcare providers. Similarly, our proposed AI-assisted tool for online art therapy promotes therapist-client interactions by involving clients so that it would ultimately improve online art therapy sessions.

6.4 Toward Designing a Collaborative AI-assisted Assessment Tool for Art Therapists

This pilot study revealed preliminary evidence that an AI-assisted assessment tool is feasible in supporting art therapists. Thus, we suggest an opportunity for a collaborative AI-assisted assessment tool that collects data from clients and delivers relevant information to therapists to support the decision-making during online therapy sessions. We found that art therapists want an AI-assisted assessment tool to collect data from clients and deliver it to them to interpret drawings. The key data from clients include how clients drew the drawings and what intentions they have for drawing certain elements. This finding indicates the importance of collaboration between AI, therapists, and clients to enhance the assessment process in online art therapy practices. To design such a collaborative AI-assisted tool, it is important to note potential privacy issues when collecting data from clients. Similar to how AI systems should ensure the protection of patients' data (Wiens et al., 2019), AI-assisted tools for art therapy should also prevent privacy issues when collecting clients' data about their behaviors, emotions, or thoughts about their drawings. Moreover, a collaborative AI-assisted tool for online art therapy should enable art therapists to obtain reliable information that describes clients' drawings. Our finding shows that art therapists expect AI-provided information should be derived from reliable resources, including the data (e.g., clients' behaviors and emotions when they draw) collected from clients both actively and passively. This finding also suggests a definition of AI accountability from art therapists' perspectives. Moltubakk Kempton and Vassilakopoulou (2021) reported that prior studies lack in providing a clear

meaning of AI accountability in healthcare. Aligning with this finding, our pilot study presents the necessity of future studies to better understand and improve the accountability of AI-assisted tools for art therapists. In addition, it is important for an AI-assisted tool to provide relevant information at the right time without adding burdens to the current practices of art therapists. Yang et al. (2019) pointed out that AI-based clinical decision tools often failed to be applicable in clinical practice because of a lack of integrity in decision-making processes in a medical context. Thus, Yang et al. (2019) also suggested that AI systems for supporting clinical decisions should be “unremarkable” that should intervene in a relevant time and place when physicians need AI support. Similarly, since our pilot study presented the concept of unremarkable AI can be applied to the art therapy context without disrupting therapists’ own practices, potential AI systems would need to provide art therapists with relevant objective information about clients’ drawings. This approach may enhance the AI-therapist collaboration in the assessment process during online therapy sessions. Therefore, we extend the existing discussions about privacy issues of data to be shared with AI systems and accountability of AI systems in healthcare to online art therapy sessions. We also propose potential directions of future research to design a collaborative AI-assisted assessment tool that facilitates information sharing between therapists and clients in online therapy sessions.

7 Limitations and Future Work

While we presented the opportunity for AI-assisted tools to support online art therapy practices, our study has several limitations. First, while multiple drawing techniques (e.g., Mandala drawing, “Person Picking an Apple from a Tree” drawing) were used by art therapists, we only focused on the use of an AI tool for conducting the HTP test in online settings. The validity of interpretations based on the projective drawings from most techniques including the HTP test in art therapy tends to be unclear as it mostly involves therapists’ subjective interpretations of clients’ drawings. Nevertheless, since the HTP test has a general guide from Buck’s manual (Buck, 1948) that enables our proposed prototype to provide therapists with information about clients’ drawings, we chose the HTP test for our prototype development. Second, due to the small sample size, the findings may be specific to the particular participants that we conducted interviews with. Although we recruited diverse participants in different age and work experience, further research with larger sample size is necessary to generalize our findings. Third, even though the Wizard of Oz approach was useful for a quick implementation and simulate a fully functional AI-assisted tool, it has some limitations such as restricted features of the prototype. Particularly, there still remain questions about the validity of our prototype in a real-life setting. Our participants may have different experiences regarding using our prototype in the actual online sessions with their clients. Despite of such limitations of Wizard of Oz approach, our study provides the CSCW community with insights about designing an AI-assisted tool

for art therapists in online settings. Similar to prior studies that used the Wizard of Oz approach to understand user perceptions toward AI (Browne, 2019; Matic et al., 2017), our study presented preliminary findings on usability of an AI-assisted assessment tool for online therapy sessions and potential concerns of art therapists about the tool. Particularly, our findings may be applicable to other online healthcare contexts (e.g., remote consultations) since many healthcare providers may experience similar challenges as our participants did. Hence, future studies remain to examine AI-assisted tools for various assessment techniques in real-world online therapy practices in different healthcare contexts, including art therapy sessions.

8 Conclusions

The goal of this pilot study was to develop prototype of an AI-assisted assessment tool that supports art therapists in online therapy sessions and evaluate its usability. This study contributes insights for developing AI-assisted online assessment tools to the CSCW community by extending empirical understanding of art therapists' perceptions toward an AI-assisted prototype. Adopting the Wizard of Oz approach with 10 art therapists, our study revealed unique needs and concerns of art therapists when they interact with our prototype. Based on our findings, we encourage researchers in the CSCW community to investigate the feasibility and validity of AI-assisted tools to support diverse stakeholders including art therapists, physicians, psychiatrists, and clients.

Acknowledgments

We would like to thank art therapists who participated in our study. We also appreciate HCAIL members for their constructive feedback on our initial manuscript. This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2020R1G1A1009133).

References

- AlSadoun, W., N. Alwahaibi, and L. Altwayan (2020): 'Towards Intelligent Technology in Art Therapy Contexts'. In: *International Conference on Human-Computer Interaction*. pp. 397–405.
- Amershi, S., D. Weld, M. Vorvoreanu, A. Fourney, B. Nushi, P. Collisson, J. Suh, S. Iqbal, P. N. Bennett, K. Inkpen, et al. (2019): 'Guidelines for human-AI interaction'. In: *Proceedings of the 2019 chi conference on human factors in computing systems*. pp. 1–13.

- Atkins, M. (2007): 'Using digital photography to record clients' art work'. *International Journal of Art Therapy*, vol. 12, no. 2, pp. 79–87.
- Austin, B. D. (2009): 'Renewing the debate: Digital technology in art therapy and the creative process'. *Art Therapy*, vol. 26, no. 2, pp. 83–85.
- Bhat, K. S., M. Jain, and N. Kumar (2021): 'Infrastructuring Telehealth in (In) Formal Patient-Doctor Contexts'. *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW2, pp. 1–28.
- Blease, C., A. Kharko, M. Annoni, J. Gaab, and C. Locher (2021): 'Machine Learning in Clinical Psychology and Psychotherapy Education: A Mixed Methods Pilot Survey of Postgraduate Students at a Swiss University'. *Frontiers in public health*, vol. 9, pp. 273.
- Browne, J. T. (2019): 'Wizard of Oz Prototyping for Machine Learning Experiences'. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA, p. 1–6, Association for Computing Machinery.
- Buçinca, Z., M. B. Malaya, and K. Z. Gajos (2021): 'To trust or to think: cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making'. *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW1, pp. 1–21.
- Buck, J. N. (1948): 'The HTP technique; a qualitative and quantitative scoring manual.'. *Journal of clinical psychology*.
- Buck, J. N. (1950): 'The use of the House-Tree-Person test in a case of marital discord'. *Journal of projective techniques*, vol. 14, no. 4, pp. 405–434.
- Caparros-Gonzalez, R. A., A. de la Torre-Luque, C. Diaz-Piedra, F. J. Vico, G. Buela-Casal, D. Dowling, and S. Thibeau (2018): 'Listening to relaxing music improves physiological responses in premature infants'. *Advances in Neonatal Care*, vol. 18, no. 1, pp. 58–69.
- Carney, M., B. Webster, I. Alvarado, K. Phillips, N. Howell, J. Griffith, J. Jongejan, A. Pitaru, and A. Chen (2020): 'Teachable machine: Approachable Web-based tool for exploring machine learning classification'. In: *Extended abstracts of the 2020 CHI conference on human factors in computing systems*. pp. 1–8.
- Chou, T. (2021): 'Apply explainable AI to sustain the assessment of learning effectiveness'. In: *12th International Multi-Conference on Complexity, Informatics and Cybernetics, IMCIC 2021*. pp. 113–118.
- Choudhry, R. and C. Keane (2020): 'Art Therapy During A Mental Health Crisis: Coronavirus Pandemic Impact Report'.
- Chromik, M. (2020): 'reshape: A framework for interactive explanations in xai based on shap'. In: *Proceedings of 18th European Conference on Computer-Supported Cooperative Work*.
- Coghlan, S., T. Miller, and J. Paterson (2021): 'Good Proctor or "Big Brother"? Ethics of Online Exam Supervision Technologies'. *Philosophy & Technology*, pp. 1–26.
- Cohen, J., J. Wright-Berryman, L. Rohlfs, D. Wright, M. Campbell, D. Gingrich, D. Santel, and J. Pestian (2020): 'A feasibility study using a machine learning suicide risk prediction model based on open-ended interview language in adolescent therapy sessions'. *International journal of environmental research and public health*, vol. 17, no. 21, pp. 8187.
- Collie, K. and D. Cubranić (2002): 'Computer-supported distance art therapy: A focus on traumatic illness'. *Journal of Technology in Human Services*, vol. 20, no. 1-2, pp. 155–171.

- Collie, K., D. Cubranic, and K. S. Booth (1998): 'Participatory design of a system for computer-supported distance art therapy'. In: *Proceedings of the participatory design conference*. pp. 29–36.
- Collie, K., S. Prins Hankinson, M. Norton, C. Dunlop, M. Mooney, G. Miller, and J. Giese-Davis (2017): 'Online art therapy groups for young adults with cancer'. *Arts & Health*, vol. 9, no. 1, pp. 1–13.
- Dahlbäck, N., A. Jönsson, and L. Ahrenberg (1993): 'Wizard of oz studies—why and how'. In: *Proceedings of the 1993 International Workshop on Intelligent User Interfaces: January*. pp. 4–7.
- Datlen, G. W. and C. Pandolfi (2020): 'Developing an online art therapy group for learning disabled young adults using WhatsApp'. *International Journal of Art Therapy*, vol. 25, no. 4, pp. 192–201.
- Denecke, K., S. L. Hochreutener, A. Pöpel, and R. May (2018a): 'Self-anamnesis with a conversational user interface: concept and usability study'. *Methods of information in medicine*, vol. 57, no. 05/06, pp. 243–252.
- Denecke, K., S. Lutz Hochreutener, A. Pöpel, and R. May (2018b): 'Talking to ana: A mobile self-anamnesis application with conversational user interface'. In: *Proceedings of the 2018 international conference on digital health*. pp. 85–89.
- Diggs, L. A., M. Lubas, and G. De Leo (2015): 'Use of technology and software applications for therapeutic collage making'. *International Journal of Art Therapy*, vol. 20, no. 1, pp. 2–13.
- Ewbank, M., R. Cummins, V. Tablan, A. Catarino, S. Buchholz, and A. Blackwell (2021): 'Understanding the relationship between patient language and outcomes in internet-enabled cognitive behavioural therapy: A deep learning approach to automatic coding of session transcripts'. *Psychotherapy Research*, vol. 31, no. 3, pp. 300–312.
- Frick, N. R., H. L. Möllmann, M. Mirbabaie, and S. Stieglitz (2021): 'Driving digital transformation during a pandemic: Case study of virtual collaboration in a German Hospital'. *JMIR medical informatics*, vol. 9, no. 2, pp. e25183.
- Goldberg, S. B., M. Tanana, Z. E. Imel, D. C. Atkins, C. E. Hill, and T. Anderson (2021): 'Can a computer detect interpersonal skills? Using machine learning to scale up the Facilitative Interpersonal Skills task'. *Psychotherapy Research*, vol. 31, no. 3, pp. 281–288.
- Gómez Penedo, J. M., B. Schwartz, J. Giesemann, J. A. Rubel, A.-K. Deisenhofer, and W. Lutz (2021): 'For whom should psychotherapy focus on problem coping? A machine learning algorithm for treatment personalization'. *Psychotherapy Research*, pp. 1–14.
- Groth-Marnat, G. and L. Roberts (1998): 'Human figure drawings and house tree person drawings as indicators of self-esteem: A quantitative approach'. *Journal of clinical psychology*, vol. 54, no. 2, pp. 219–222.
- Hacmun, I., D. Regev, and R. Salomon (2018): 'The principles of art therapy in virtual reality'. *Frontiers in psychology*, vol. 9, pp. 2082.
- Hamdy, H., J. Sreedharan, J. I. Rotgans, N. Zary, S. A. Bahous, M. Venkatramana, E. AbdelFattah Elzayat, P. Lamba, S. K. Sebastian, and N. K. A. Momen (2021): 'Virtual Clinical Encounter Examination (VICEE): A novel approach for assessing medical students' non-psychomotor clinical competency'. *Medical Teacher*, pp. 1–7.

- Holtzblatt, K. and H. Beyer (1997): *Contextual design: defining customer-centered systems*. Elsevier.
- Jacobs, M., J. He, M. F. Pradier, B. Lam, A. C. Ahn, T. H. McCoy, R. H. Perlis, F. Doshi-Velez, and K. Z. Gajos (2021): 'Designing AI for Trust and Collaboration in Time-Constrained Medical Decisions: A Sociotechnical Lens'. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. pp. 1–14.
- Johnson, E. M., M. F. Hamilton, R. S. Watson, R. Claxton, M. Barnett, A. E. Thompson, and R. Arnold (2017): 'An Intensive, Simulation-Based Communication Course for Pediatric Critical Care Medicine (PCCM) Fellows'. *Pediatric critical care medicine: a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies*, vol. 18, no. 8, pp. e348.
- Jones, B., S. Prins Hankinson, K. Collie, and A. Tang (2014): 'Supporting non-verbal visual communication in online group art therapy'. In: *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. pp. 1759–1764.
- Kim, S.-i. (2008): 'Computer judgment of main color in a drawing for art psychotherapy assessment'. *The Arts in Psychotherapy*, vol. 35, no. 2, pp. 140–150.
- Kim, S.-i., J. Bae, and Y. Lee (2007): 'A computer system to rate the color-related formal elements in art therapy assessments'. *The Arts in Psychotherapy*, vol. 34, no. 3, pp. 223–237.
- Kim, S.-i. and I. A. Hameed (2009): 'A computer system to rate the variety of color in drawings'. *Art Therapy*, vol. 26, no. 2, pp. 73–79.
- Kim, S.-i., H.-S. Kang, and K.-E. Kim (2008): 'Computer determination of placement in a drawing for art therapy assessments'. *The Arts in Psychotherapy*, vol. 35, no. 1, pp. 49–59.
- Kim, S.-i., H.-S. Kang, and Y.-H. Kim (2009): 'A computer system for art therapy assessment of elements in structured mandala'. *The Arts in Psychotherapy*, vol. 36, no. 1, pp. 19–28.
- Kim, T., Y. Yoon, K. Lee, K. Kwahk, and N. Kim (2021): 'Application of Deep Learning in Art Therapy'. *International Journal of Machine Learning and Computing*, vol. 11, no. 6.
- Lee, M. H., D. P. Siewiorek, A. Smailagic, A. Bernardino, and S. B. Bermúdez i Badia (2021): 'A Human-AI Collaborative Approach for Clinical Decision Making on Rehabilitation Assessment'. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. pp. 1–14.
- Levy, C. E., H. Spooner, J. B. Lee, J. Sonke, K. Myers, and E. Snow (2018): 'Telehealth-based creative arts therapy: Transforming mental health and rehabilitation care for rural veterans'. *The Arts in Psychotherapy*, vol. 57, pp. 20–26.
- Lewis, J. R. (1991): 'Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ'. *ACM Sigchi Bulletin*, vol. 23, no. 1, pp. 78–81.
- Li, C.-Y., T.-J. Chen, C. Helfrich, and A.-W. Pan (2011): 'The development of a scoring system for the Kinetic House-Tree-Person drawing test'. *Hong Kong Journal of Occupational Therapy*, vol. 21, no. 2, pp. 72–79.
- Lloyd, B. and M. Usiskin (2020): 'Reimagining an emergency space: practice innovation within a frontline art therapy project on the France-UK border at Calais'. *International Journal of Art Therapy*, vol. 25, no. 3, pp. 132–142.

- Matic, A., M. Pielot, and N. Oliver (2017): "OMG! How Did It Know That?": Reactions to Highly-Personalized Ads'. In: *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization*. New York, NY, USA, p. 41–46, Association for Computing Machinery.
- Miner, A. S., N. Shah, K. D. Bullock, B. A. Arnow, J. Bailenson, and J. Hancock (2019): 'Key considerations for incorporating conversational AI in psychotherapy'. *Frontiers in psychiatry*, vol. 10, pp. 746.
- Moltubakk Kempton, A. and P. Vassilakopoulou (2021): 'Accountability, Transparency and Explainability in AI for Healthcare'. In: *Infrahealth 2021-Proceedings of the 8th International Conference on Infrastructures in Healthcare 2019*.
- Parker-Bell, B. (1999): 'Embracing a future with computers and art therapy'. *Art Therapy*, vol. 16, no. 4, pp. 180–185.
- Pearson, R., D. Pisner, B. Meyer, J. Shumake, and C. G. Beevers (2019): 'A machine learning ensemble to predict treatment outcomes following an Internet intervention for depression'. *Psychological medicine*, vol. 49, no. 14, pp. 2330–2341.
- Power, N., R. Dolby, and D. Thorne (2021): "Reflecting or frozen?"The impact of Covid-19 on art therapists working with people with a learning disability'. *International Journal of Art Therapy*, pp. 1–12.
- Raj, M., B. Iott, D. Anthony, and J. Platt (2022): 'Family Caregivers' Experiences With Telehealth During COVID-19: Insights From Michigan'. *The Annals of Family Medicine*, vol. 20, no. 1, pp. 69–71.
- Rohail, I. (2015): 'Underlying Risk Factors of Personality of Clinical Patients Illustrated Through House Tree Person Test'. *Academic Research International*, vol. 6, no. 4, pp. 193–203.
- Romanovskyi, O. and N. P. A. Knysh (2021): 'Elomia Chatbot: the Effectiveness of Artificial Intelligence in the Fight for Mental Health'.
- Schwartz, B., Z. D. Cohen, J. A. Rubel, D. Zimmermann, W. W. Wittmann, and W. Lutz (2021): 'Personalized treatment selection in routine care: Integrating machine learning and statistical algorithms to recommend cognitive behavioral or psychodynamic therapy'. *Psychotherapy Research*, vol. 31, no. 1, pp. 33–51.
- Sebri, V., S. F. M. Pizzoli, L. Savioni, and S. Triberti (2020): 'Artificial Intelligence in mental health: professionals' attitudes towards AI as a psychotherapist'. *Annual Review of Cybertherapy and Telemedicine 2020*, pp. 229.
- Shorey, S., E. Ang, J. Yap, E. D. Ng, S. T. Lau, and C. K. Chui (2019): 'A virtual counseling application using artificial intelligence for communication skills training in nursing education: development study'. *Journal of medical Internet research*, vol. 21, no. 10, pp. e14658.
- Snyder, K. (2021): 'The digital art therapy frame: creating a 'magic circle'in teletherapy'. *International Journal of Art Therapy*, pp. 1–7.
- Solomonov, N., J. Lee, S. Banerjee, C. Flückiger, D. Kanellopoulos, F. M. Gunning, J. A. Sirey, C. Liston, P. J. Raue, T. D. Hull, et al. (2020): 'Modifiable predictors of nonresponse to psychotherapies for late-life depression with executive dysfunction: a machine learning approach'. *Molecular Psychiatry*, pp. 1–9.
- Strauss, A. and J. M. Corbin (1997): *Grounded theory in practice*. Sage.

- Su, Y.-H., M.-J. Su, Y.-T. Chen, H.-W. Zhang, R. Chen, S. Lai, Y.-S. Lin, S.-J. Chen, and H.-S. Chen (2009): 'Digital art and quality living for elderly people and terminal cancer patients'. In: *2009 1st International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology*. pp. 59–62.
- Wiens, J., S. Saria, M. Sendak, M. Ghassemi, V. X. Liu, F. Doshi-Velez, K. Jung, K. Heller, D. Kale, M. Saeed, et al. (2019): 'Do no harm: a roadmap for responsible machine learning for health care'. *Nature medicine*, vol. 25, no. 9, pp. 1337–1340.
- Xie, Y., M. Chen, D. Kao, G. Gao, and X. Chen (2020): 'CheXplain: Enabling Physicians to Explore and Understand Data-Driven, AI-Enabled Medical Imaging Analysis'. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. pp. 1–13.
- Yang, Q., A. Steinfeld, and J. Zimmerman (2019): 'Unremarkable ai: Fitting intelligent decision support into critical, clinical decision-making processes'. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. pp. 1–11.
- Yao, L., X. Zhao, Z. Xu, Y. Chen, L. Liu, Q. Feng, and F. Chen (2020): 'Influencing Factors and Machine Learning-Based Prediction of Side Effects in Psychotherapy'. *Frontiers in Psychiatry*, vol. 11.
- Yu, Y. Z., C. Yu Ming, M. Yue, J. Hai Li, and L. Ling (2016): 'House–Tree–Person drawing therapy as an intervention for prisoners' prerelease anxiety'. *Social Behavior and Personality: an international journal*, vol. 44, no. 6, pp. 987–1004.
- Zubala, A. and S. Hackett (2020): 'Online art therapy practice and client safety: a UK-wide survey in times of COVID-19'. *International Journal of Art Therapy*, vol. 25, no. 4, pp. 161–171.

Aisha Ahmed AlArfaj & Ellis Solaiman (2022): Comparing different qualitative methods to understand user experience in Saudi Arabia. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_ep05

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_ep05

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

Comparing different qualitative methods to understand user experience in Saudi Arabia

Aisha Ahmed AlArfaj & Ellis Solaiman

College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia

School of Computing, Newcastle University, Newcastle Upon Tyne, UK

aiaalarfaj@pnu.edu.sa

ellis.solaiman@newcastle.ac.uk

Abstract. The HCI field has seen a growing body of qualitative research, making use of a wide range of activities and methods. Interviews and workshops are some of the main techniques used to help understand user needs and to conduct co-design activities with them. However, these methods might be conducted in various ways and have different advantage and disadvantages. An important aspect influencing the types of activities and methods used is the culture of research participants. This paper aims to compare the research methods conducted in the context of the Saudi Arabian culture. It provides a reflection on the methods used to understand user needs when designing social commerce platforms, including interviews, co-design workshops and critique design workshops. We found that each method has its positives and negatives in terms of user preferences, and can help to obtain useful information at different levels of detail. For example, conducting semi-structured interviews by text was preferred by participants who are at home with their families. However, they can be slower than other methods.

Introduction

This study employed a combination of a user-centred approach and co-design methodology to develop an in-depth understanding of the current use of social commerce (s-commerce) platforms in Saudi Arabia, aiming to improve user experience and enhance trust in these platforms. The aim of this paper is to present and discuss the methods conducted in this research, highlighting their advantages and disadvantages from practical and cultural perspectives.

A qualitative approach is needed to help researchers study and understand user needs. Within this approach, several methods can be used in human-computer interaction (HCI) research. However, some of these methods present potential issues or challenges, such as high financial cost or being very time-consuming to implement. Moreover, the same methods can be conducted in various ways to obtain different information and details. For example, interviews could be conducted by phone, message or face-to-face. Moreover, they could include activities such as sharing maps and picture cards, or be combined with other methods, such as diary studies, to collect more information on participants' experiences.

For this research, five studies were conducted, employing a mix of methods (interviews, observation, co-design workshops and critique workshops). The first study sought to gain a general understanding of the use of social commerce based on buyer experiences (using text and voice interviews) (AlArfaj et al., 2019). The second study developed a broad knowledge of the use of social commerce from the sellers' perspective, based on understanding what sellers provide to buyers and how they attract them and gain their trust (using face-to-face and text interviews with online observation)(AlArfaj and Solaiman, 2019). The third study examined sharing activities and compared the use of social commerce and e-commerce (using face-to-face interviews, including drawing a sharing map with a diary). The fourth study employed co-design workshops to finalise the design recommendations (using co-design workshops including screenshots of interfaces and card sorting) (AlArfaj and Solaiman, 2021). Finally, the fifth study aimed to provide a critique of the social commerce platform mock-ups (using critique design workshops, including the mock-ups) (AlArfaj and Solaiman, 2021).

The motivations behind this work are that there is a need to understand the routines and behaviours of s-commerce shoppers to be able to develop or improve social commerce platforms effectively. However, there are few studies that have explored the use of social media to conduct commercial activities (Huang and Benyoucef, 2013). In addition, while there is a growing body of literature on s-commerce around the world (Baethge et al., 2016), few studies have been conducted in the context of Saudi Arabia. Most of these studies employing quantitative approaches. Therefore, this novel work was conducted in a culture little explored previously, Saudi Arabia, using qualitative methods.

Background

Research approach

This research focuses on the use of s-commerce in Saudi Arabia. Social commerce can be defined as "the use of Internet-based media to enable users to participate in the selling, buying, comparing, and sharing of information about products and services in online marketplaces and communities" (Busalim et al., 2016). Most of the definitions refer to two main components: social media and e-commerce. According to recent literature (Huang and Benyoucef, 2013), there are two main categories of s-commerce:

1. Social media with commercial features enables users to have a personal profile and share information and to search for and buy products using social networking sites (SNSs).
2. Traditional e-commerce websites with social communication and sharing activities, enabling users to rate, review and post comments, as well as to join a community and be able to use social plugin and buy.

Most existing studies in the field of s-commerce have been quantitative in nature (Esmaeili and Hashemi G, 2019) and questionnaires have most commonly been used to collect data in Saudi Arabia (Alghamdi et al., 2015; Nassir and Leong, 2017). However, there is a need to gain an in-depth understanding of the current use of s-commerce and how to improve and enhance trust and for such research concerning the use of technologies and the challenges faced, it is recommended to conduct qualitative inquiries to examine the interaction behaviour (Lazar et al., 2017). Whereas quantitative studies address what people do, qualitative studies can help researchers understand why they do what they do (Dimond et al., 2012).

User-centred design (UCD) considers the user as a subject (Sanders and Stappers, 2008), which means the focus is on end users (Baxter et al., 2015). The user-centred approach can help to gather information and understand users' experiences through a variety of methods (Baxter et al., 2015). The development of the research design is based on the users' tasks and goals, investigating the context and behaviour of users and considering the users' characteristics (Rumpf, 1990). In HCI, several methods can be employed to study the use of an interface or application, the most common of which are observation, field studies, surveys (usability studies), interviews, focus groups and controlled experiments (Lazar et al., 2017).

The participatory approach considers the user a partner who is included in a co-design process. This means that people who are not designers work together creatively during the design development process (Sanders and Stappers, 2008). The results of a UCD approach should be a well-designed system that is useful and easy to use (Rumpf, 1990). UCD helps to collect and analyse user requirements, which are the features/attributes that a product/system should include or how the product/system should perform from the users' perspective (Baxter et al., 2015).

User experience is human-centred and the goal of examining the user experience is to elicit users' requirements or to evaluate an existing technology (Baxter et al., 2015). It is essential to capture the users' experiences and to understand their needs.

Various methods can be implemented to help researchers understand users, such as interviews, diary studies, surveys, card sorting, focus groups, field studies and evaluations (Baxter et al., 2015). In this paper, the focus is on exploring a range of methods: semi-structured interviews, sharing maps, diary studies, picture cards, online observation (online ethnography), co-design workshops and critique design workshops.

Qualitative interviews aim to explore and gain an in-depth understanding of a topic (Rumpf, 1990). Interviews are considered an invaluable tool in HCI research in helping the researcher to understand users' needs, practices, attitudes and behaviours (Baxter et al., 2015; Dimond et al., 2012; Maeng et al., 2016; Voidsa et al., 2004). In particular, semi-structured interviews are conducted when researchers want to learn more about specific topics; these are based on a basic script with questions to be asked, but then probe the participants to provide more information that might be relevant (Rumpf, 1990).

There are five forms of interviews: face-to-face, video with audio, audio, interactive text (e.g. online chat) and non-interactive text (e.g. email) (Baxter et al., 2015; Maeng et al., 2016). Researchers can choose the type most appropriate to their needs (Baxter et al., 2015). These types fall under two main categories, in person and mediated (remote) (Baxter et al., 2015; Martin et al., 2012). The form of the interview can affect different aspects of the process, such as expectations of attention, timing, transcription, sharing of multimedia and conversation style (Maeng et al., 2016; Voidsa et al., 2004).

Face-to-face interviews can be costly in terms of finances and time; also, there is burdens on both the interviewer and the interviewee and organising the data is a lengthy process (Maeng et al., 2016). Several studies have conducted interviews through calls (e.g. phone, Skype, etc.) to help identify practices and concepts related to the research focus or to provide an overall understanding of technology with additional, deeper insights (Al-Dawood et al., 2017; Alghamdi et al., 2015; Nylander and Rudström, 2011). However, such modes of interviewing can also be costly and create burdens on both the participants and researcher (Maeng et al., 2016).

Other forms of interview (e.g. online messages) can be beneficial for interviewing participants who are not within driving distance of the researcher or interviewer (Hillman et al., 2014), or to include participants from different geographical regions (Alsheikh et al., 2011; Al-Saggaf et al., 2004). Mediated forms are especially helpful in conducting interviews with the latter (Rumpf, 1990). Instant messages help the researcher to communicate with participants and share ideas even at a great distance (Voidsa et al., 2004). Participants might multitask when engaged in an interview via online messaging, e.g. sending messages to others or doing something else (Voidsa et al., 2004). Moreover, conducting interviews using text can save the time the researcher might otherwise

have to spend on transcription (Maeng et al., 2016; Voids et al., 2004). Therefore, mediated interviews are considered an easier option than in-person interviews (Baxter et al., 2015). However, in face-to-face interviews, there is the assurance that the participants can focus fully on the interview (Voids et al., 2004).

A previous study used quantitative analysis, focusing on word count, and qualitative codes to compare three forms of interviews, i.e. instant messaging, email and phone (Dimond et al., 2012). Although the number of words in the phone interviews was four times greater, there were no significant differences in the numbers of codes (Dimond et al., 2012). However, another study that compared phone use, instant messaging and mobile instant messaging found more data were collected using phone interviews (Maeng et al., 2016).

As already noted, interviews can either be conducted as a solo activity or in combination with another, such as card sorting or observation (Baxter et al., 2015). Integrating other methods, such as picture card sorting, can make the interviews more productive (Martin et al., 2012). Using picture cards is an artefact-based method, in which the researcher use images and words to help participants recall and recount actual experiences and details (Martin et al., 2012). Asking users in an interview how they use a system may not yield accurate answers (Lazar et al., 2017). However, giving the participants picture cards of the interfaces can gain more valid details because they help users recall their experiences and stories and they can represent their wishes (Martin et al., 2012). This method is more productive than others, supporting users' recall and helping them become more involved in the conversation.

The sharing map is one of the methods that can be used in interviews to understand participants' interactions with their circle with regard to online shopping (Hillman et al., 2013). A sharing map can show both the social networks participants use and – perhaps more importantly – the significant people with whom they share their shopping experiences and information.

There also methods that can be conducted alongside interviews, such as observation and diary studies (Baxter et al., 2015). Observational methods have been used, for example, as an additional means of gaining an overall understanding of interactions within the platform (Chen, 2010; Evans et al., 2018; Moser et al., 2017). Researchers might observe the activities and interactions within an s-commerce platform to enrich the data collected (Dye et al., 2016). Observation can also help collect baseline information (Martin et al., 2012). Moreover, it can be used as a means of online ethnography, enabling the researcher to gather insights from online communities (Tomitsch et al., 2018). Online ethnography can be conducted to investigate issues related to the online community (Gheitasy et al., 2015). This method helps to study online interactions among those participating in an online platform (Tomitsch et al., 2018).

Diary studies can be used to collect data across time, including participants' thoughts, feelings and behaviours over a specific period (Martin et al., 2012). Diary studies are usually conducted with other methods, such as interviews (Baxter et al., 2015). Moreover, digital diaries can be used, enabling participants to submit their

diaries online and include photos and text entries (Martin et al., 2012). The benefits of this method are that it can help researchers enrich the data collected without the need to be present (Baxter et al., 2015).

The other approach is participatory design (PD), which can be defined as "a set of theories, practices, and studies related to end-users as full participants in activities leading to software and hardware computer products and computer-based activities" (Muller and Druin, 2012). PD can also be called co-creation/co-design (Sanders and Stappers, 2008). In PD methods, all the participants contribute to discussions during workshops (Dillahunt and Malone, 2015). Co-design has been growing in the PD field (Sanders and Stappers, 2008); co-design workshops help to ensure that the users' needs are addressed during the design process, the application being designed with them rather than for them (Tomitsch et al., 2018). The co-design process can help provide users with a higher quality service experience that better meets their needs (Steen et al., 2011). Co-design workshops can be used in any stage of the design process (Tomitsch et al., 2018). They can help designers to generate better ideas and acquire greater knowledge about customers' needs (Steen et al., 2011). They can be used in the research phase to help gain a complete overview of the situation (Tomitsch et al., 2018). They can also be used during the prototype phase to iterate concepts rapidly (Tomitsch et al., 2018).

An early method for evaluate design is to conduct design critique workshops, which can help to reflect on the design rather than testing it (Greenberg and Buxton, 2008). Design critique focuses on evaluating existing ideas, not developing new ideas (Tomitsch et al., 2018). Such evaluation can be undertaken by presenting the design to the participants and asking them to identify what they like and dislike (Greenberg and Buxton, 2008; Tomitsch et al., 2018).

Qualitative and design studies in social (s-)commerce

The majority of existing studies in s-commerce have been quantitative and have employed surveys (Esmaeili and Hashemi G, 2019; Han et al., 2018; Zhang and Benyoucef, 2016). It is recommended that future studies in s-commerce conduct qualitative research to obtain more empirical evidence with respect to consumer behaviour (Han et al., 2018) as few studies have used qualitative methods to understand users' behaviours and needs with regard to s-commerce (Moser et al., 2017; Shin, 2013).

A previous study explored the use of Facebook as an s-commerce platform to trade products within a mom-to-mom buy and sell group (Moser et al., 2017), with the mothers as both buyers and sellers. The study entailed semi-structured interviews and observation of the group for five months to collect field notes of the members' interactions. Observation helped contribute to a general understanding of how Facebook buy and sell groups work. Another study investigated the use of Facebook to conduct transactions (Evans et al., 2018). The methods used were semi-structured interviews and monitoring of Facebook bartering groups for eight months, which provided an in-depth understanding of the interactions between members. Another study was also conducted with Facebook to identify the factors

that affect consumers' intentions (Anari et al., 2014). Interviews were undertaken to identify new factors that can affect consumers' intentions and the reasons why.

In a study that used both semi-structured interviews and sharing maps to capture personal data on shopping and relations (Hillman et al., 2013), the use of group shopping and friendship networks was investigated to propose the optimal design for s-commerce based on the consumers' experiences with a view to improving the user experience. Another study on s-commerce examined the impact of social information on consumers' decision processes (Chen, 2010). Semi-structured interviews were conducted after the participants performed a task using two sites. The data provided in-depth information regarding the motivation behind consumers' choice of product. Moreover, observing the participants while they performed the tasks provided information concerning social features and their role in buying a product. Furthermore, a previous study explored the role of s-commerce in enhancing consumers' trust (Ji et al., 2020). Again, semi-structured interviews were conducted, as well as direct observation of the s-commerce platforms (companies' websites and WeChat). These methods helped identify issues related to consumers' trust.

A previous study that conducted mock-ups used a survey as the main research method (Suraworachet et al., 2012). The survey employed mock-ups, incorporating a scenario and a rating scale, then asked for any additional suggestions. It evaluated the effect of s-commerce features on intention to buy over Facebook. Another study used experiment-based surveys to explore the influence of s-commerce features on users' attitudes towards a manufactured e-commerce website (Friedrich et al., 2016). The e-commerce website was created based on previous literature. The researchers provided design implications based on their results.

Although these studies investigated the use of s-commerce and consumers' behaviours, they did not include sufficient details about the role of the features and design in enhancing consumers' trust. Moreover, most of the studies focused on Facebook, WeChat and various websites. As far as can be ascertained, there are no extant studies on s-commerce that have used co-design or critique workshops. There are only limited studies that have discussed the design of s-commerce platforms, the features necessary and their impact.

Sample size

The recommended sample size for interviews examining users' perceptions varies. One study recommended having at least six participants to gather users' perceptions (Huang et al., 2012). However, based on expert opinion, another argued that the optimal number of participants depends on the method, i.e. 12–20 participants for interviews and 4–12 participants per group for focus groups, with 3–4 groups in all (Baxter et al., 2015). Previous studies that have been conducted to understand users' experiences and report their perceptions have recruited 14–18 participants (Al-Dawood et al., 2017, 2021; Al-Saggaf et al., 2004; Hillman et al., 2013; Nassir

and Leong, 2018). For the co-design workshops and design critique workshops in this study, it was recommended to have 3–10 participants (Tomitsch et al., 2018).

Culture

The main challenges that may face researchers in Saudi Arabia are participants' privacy concerns when talking to a stranger and cross-gender communication (Abokhodair and Vieweg, 2016; Nassir and Leong, 2017). It is difficult to recruit research participants in Saudi Arabia (Abokhodair and Vieweg, 2016), especially due to privacy concerns, which can be significant (Nassir and Leong, 2017). To make it easier to secure a sample, the snowball sampling approach should be used, as recommended by others (Huang and Benyoucef, 2013), to help participants trust the researcher to a certain extent (Nassir and Leong, 2017). Using this means, the researcher recruits participants who are recommended by personal and professional connections and they in turn recruit others (Abokhodair and Vieweg, 2016). However, this approach may not result in participants with diverse backgrounds (Al-Dawood et al., 2017, 2021). Therefore, it is useful also to use another approach, for example by sending invitations through social networks (Alghamdi et al., 2015) (including an online sign-up form (Al-Dawood et al., 2021)), mixing approaches to reach a wider range of participants from diverse backgrounds (Al-Dawood et al., 2017, 2021). Recruiting participants by email has been found not to be as effective as instant messaging (Al-Saggaf et al., 2004). It is also important to describe the data collection method(s) in the invitation letter, particularly as some methods, such as focus groups, are not commonly used in Saudi Arabia (Alghamdi et al., 2015). Indeed, although Saudi Arabia is a modern society, it is conservative (Al-Saggaf, 2011) and therefore great consideration should be given to the data collection methods used when conducting studies in studies, such as interviews (Al-Dawood et al., 2017) (Al-Saggaf, 2011).

In cross-gender communication, entailing interaction between females and males, the women should be chaperoned (Alsheikh et al., 2011). Moreover, a male researcher may ask female participants to gain permission from their guardians prior to conducting interviews (Al-Dawood et al., 2017) (Al-Saggaf, 2011). It has been considered essential not to initiate direct contact with females, but to draw on male relatives to mediate talks (Nassir and Leong, 2018). However, the law of guardianship changed in 2019 to allow women above 21 years of age to travel without permission and to apply for a passport (Krimly, 2019). Therefore, asking the guardian's permission might not be necessary from the legal perspective, but the cultural norms persist (Al-Dawood et al., 2021).

For chaperoned interviews, the researcher can use probes to collect data supplementing the initial responses as the answers from participants might be affected by the presence of the chaperones (Nassir and Leong, 2018). Moreover, online communication can be beneficial for several reasons. First, it facilitates cross-gender communication (Al-Saggaf et al., 2004). As well as being low cost, online communications have been found to be effective in gaining in-depth information about an individual's experiences (Al-Saggaf et al., 2004). Sometimes,

female participants might prefer to be interviewed anonymously via online chat, for example using Facebook Messenger, to prevent communicating with a (male) stranger (Al-Dawood et al., 2021) (Almakrami, 2015). Conducting interviews by telephone is also considered an appropriate method due to the culture of gender segregation (Alghamdi et al., 2015). Female participants can also communicate with a male researcher and pass on information through social media (Nassir and Leong, 2018). Equally, it is better for a female researcher to conduct interviews with male participants by telephone, not face to face (Flechais et al., 2013), or through social networks to avoid the need for the researcher to have a chaperone (Nassir and Leong, 2018).

For qualitative studies in Saudi Arabia, social media can be a good source of information from participants, especially in the case of cross-gender communication (Nassir and Leong, 2018). Participants may be willing to share photographs and videos with researchers using social networking (Nassir and Leong, 2018). However, conducting an initial interview before asking participants to share information over social media may help to build trust, especially as the participants know more about the researcher and vice-versa (Nassir and Leong, 2018).

Deciding where to conduct a face-to-face meeting is another challenge that faces the researcher as it is not permitted to do so in private premises and rent a meeting room is an expensive option (Alghamdi et al., 2015). It is common in Saudi Arabia for females and males to have separate spaces in places such as schools, universities and government departments (Flechais et al., 2013). Therefore, separate sessions for focus groups should be conducted for females and males in different gender-appropriate locations, the female sessions being run by a female researcher/moderator and the male sessions being run by a male researcher/moderator (Alghamdi et al., 2015).

Interviews should be conducted in a public place or workplace (Flechais et al., 2013). A public place could be a restaurant that provides family rooms (Alghamdi et al., 2015). While interviews can be undertaken in the participant's home, it can be a challenging environment because of the presence of other family members and the potential for interruptions and noise, for example from children and extended family (Nassir and Leong, 2018). It is not usual to interview participants in their homes even if they are the same gender (Alghamdi et al., 2015), but a previous study conducted in the western region of Saudi Arabia, where gender segregation is more relaxed, did so (Nassir and Leong, 2018). The researcher, who recruited participants through snowball sampling, was treated as a guest and provided with hospitality (Nassir and Leong, 2018).

In Saudi Arabia, many participants may reject compensation for participation, for example a gift card, preferring to take part voluntarily (Al-Dawood et al., 2017). It might be more appropriate to give a gift, such as chocolates or a pen, rather than cash or a gift card (Al-Dawood et al., 2021; Nassir and Leong, 2017). Indeed, compensation for Saudi participants should not be monetary.

The context of Saudi Arabia

In Saudi Arabia, s-commerce has grown rapidly (Alotaibi et al., 2019). It has been used by women entrepreneurs to present and sell their products, for example, those that are handmade or customised. Previous studies have examined entrepreneurs' use of social media applications to conduct commercial activities, such as implementing a marketing strategy (Alotaibi et al., 2019; AlGhamdi and Reilly, 2013). Such studies used questionnaires or online observation and recorded social media data through Web analytics tools. They found that there are cultural boundaries (i.e. familial, governmental and political). However, their results are still being developed. Other studies have focused on factors that are positively associated with behavioural intentions, such as social support and s-commerce constructs (Makki and Chang, 2015; Sheikh et al., 2017), as well as factors that influence the adoption of s-commerce (Abed et al., 2015). They all employed a quantitative approach based on questionnaires.

There are very few studies that have investigated factors affecting consumer trust. Examples of such studies include that of Alotaibi, Alkathlan, and Alzeer (Alotaibi et al., 2019), which looks at factors such as social media influencers (SMIs), key opinion leaders (KOLs), and consumer feedback. Another interesting work demonstrates how trust mediates the relationship between social support and s-commerce intentions (Al-Tit et al., 2020). However, in general, there is a lack of research focused on identifying the mechanisms that can enhance trust in online s-commerce through design. Moreover, the studies that have been conducted only tested a limited number of factors and solely employed a quantitative approach.

Methods

This study employed a user-centred and participatory approach to understand the use of s-commerce and how its features influence users. The research comprised five studies aimed at understanding user behaviour and designing and evaluating an s-commerce platform. The main aim was to understand the current use of s-commerce and identify how the user experience and enhance trust focusing on the features of the application. Table I shows the methods and tools used, as well as the number of participants in each study.

The first study entailed semi-structured interviews with buyers conducted by telephone and instant messaging (e.g. WhatsApp). The questions asked about their use of social media applications, the commercial use of social media applications and the use of e-commerce websites (AlArfaj et al., 2019).

The second study comprised semi-structured, face-to-face interviews with buyers (AlArfaj and Solaiman, 2019). The interview protocol had three parts: (i) asking participants to draw a sharing map of their social networks and the people with whom they shared their shopping information; (ii) asking participants to imagine conducting commercial activity through a social network and through an e-commerce website and explain the differences in use and issues of trust; (iii)

Table I: Summary of methods conducted in the study.

	Main method(s)	Tools	Number of participants
1	Semi-structured interviews with buyers	By text	18
		By voice (telephone or online such as Skype)	8
2		Face to face with activities	17
3	Semi-structured interviews and online observation with sellers	Face to face	7
		By text	10
4	Co-design workshops	Screenshots of interfaces (Instagram and Amazon) + online access to platforms	10
5	Critique workshops	System usability scale (SUS) + interactive prototype + screenshots of the interactive prototype	16

showing picture cards and asking participants what features might help them make a purchase decision.

The third study conducted semi-structured interviews with sellers to identify how they gained customers' trust. The interviews were either conducted face to face or through instant messaging. Online observation was also undertaken to examine activities and interactions.

The fourth study comprised a co-design workshop to inform on the design of an s-commerce platform and identify the features that should be included (AlArfaj and Solaiman, 2021). The workshop included screenshots of Instagram and Amazon to help participants describe what they liked on these platforms and how one might be better designed. Participants were asked to sketch a prototype, but they were unlikely to do so. Therefore, they were given cards showing the features/functions and their icons and asked to sort these cards into three groups according to importance: very important, important and not important.

Finally, the fifth study entailed the design critique workshops, presenting mock-ups including the very important and important features and asking the participants to critique the design (AlArfaj and Solaiman, 2021). In these workshops, participants were asked to identify what they liked and what they did not understand and their further suggestions for the design.

Recruitment

Three methods were used to recruit participants: sending emails, posting an invitation on social networks and snowball sampling. Posting the invitation on social networks and the snowball method were more effective than the email. A Google form was created and included in the invitation with an information sheet and description of the study. For all the studies conducted, participants were asked if they could/would take part in interviews or workshops online.

For the co-design and critique design workshops, the invitation included a message asking those happy to participate to click on the link to the Google form. The invitation also stated that I, as the researcher, would be happy to answer any questions before participants registered. For the workshops, few participants registered initially. I received several private messages asking almost identical questions about the workshop: where would it be, could they attend virtually, was it mixed gender or only female, did they need to draw something or just talk? After answering these questions directly, some of them registered. I also resent the invitation through social networks, addressing these questions in the invitation as most would not open the information or the link before knowing the basic details.

Results

The following sub-sections discuss various aspects concerning the implementation of the methods used in the different studies.

Time

The telephone interviews lasted, on average, 15 minutes. The face-to-face interviews took around 20–40 minutes. The face-to-face interviews that also included activities naturally took longer than those that did not. However, the online chats took longer, especially as sometimes the participants were busy and took 1 or 2 hours to respond, for which they apologised. Although the online interactions took longer, it was not necessary to wait to reply to participants, as I could respond immediately when I received a notification.

The face-to-face interviews were more time consuming. Travelling to the interview location took around 30 minutes as the routes in Saudi Arabia are increasingly crowded. Moreover, on arriving at the location, some of the participants did not attend, apologising that something urgent had come up. With the online messages and telephone interviews, there was no need to spend time on travelling to the location. Also, online interviews could be conducted by other means, such as mobile telephones, and they could take place at any time and in any place. Some participants interviewed through online chat sent voice notes in response instead of typing their answers, which was even better because they included more details. However, many participants preferred to type rather than sending voice messages.

In terms of the workshops, some participants who were registered did not attend, sending apologies that they had to stay with their children or they did not feel well. The workshops had to start on time as the other participants might have later appointments. To overcome the issues with last minute cancellations, I recruited the maximum number of potential participants who might attend the workshop, so that even if a few did not attend, the overall participation was reasonable.

Response: Concentration and engagement

The response, level of attention and engagement from users differed based on the form of the interview and location. Interviewing by telephone presented some challenges. First, some participants did not live in a quiet location, making it difficult to hear them or to have a smooth conversation. Some were at home with their family and children, which resulted in background noise and the participants being distracted by family members asking questions. Sometimes, the participants received another call during the interview and asked if they could just take it quickly, which was time consuming. In some cases, the participants were walking or driving during the interviews, which could be noisy and meant they were not entirely focused on the interview.

In contrast, interviewing through online messaging, for example using WhatsApp, Skype, Line, etc., helped ensure a smooth conversation without any noise. However, some participants would multitask, responding to other messages or doing something else. Before the interviews started, some participants asked if it would be alright to be "in and out", saying they would prefer this as they would be able to concentrate and answer the questions fully if, for example, they could leave and then return to a question after dealing with an interruption from a family member or background noise.

Regarding the face-to-face interviews, the attention of the participants depended on the location. The interviews held in a public place, such as coffee houses, were sometimes disrupted by noise when the café was busy. Other interviews were conducted in a meeting room at the participant's workplace or in the participant's office. In the former instance, the participants were able to concentrate fully on the interview, but in the latter case they were sometimes interrupted by colleagues calling on the telephone or dropping by to ask about something.

In terms of engagement, each form differed. In the online chat mode, it was easy to share multimedia and to ask questions freely and be open to answering questions. The participants also shared pictures of their experiences or photographs of the applications that they liked. However, with the telephone interview, it was not possible to share any images and the participants did not really ask questions; they just answered my questions and ended the interview as soon as possible. In the face-to-face interviews, I could show the participants picture cards and we could discuss them, as in online messaging. However, there were also other activities that encouraged the participants to discuss their experiences in greater depth, such as drawing a map of their social media application use and accessing their mobile devices and imagining ordering through s-commerce/ e-commerce and how they would go about it.

In terms of online chat, it was easier to follow up with participants and ask them if additional information was needed to verify their meaning. Also, the participants shared screenshots and photographs of their shopping experiences and sent them through WhatsApp with comments (e.g. that they knew the shop account from their friends or a screenshot of a recommendation from their friends or family).

In the co-design workshops, some features were not identified as important by users, but when they saw them in the mock-ups, they thought them really helpful and important. For example, the reviewer's profile was not considered important in the co-design workshops, but in the critique design workshop, they checked the profile to ensure it was not fake and viewed it as helpful.

Culture

There were some challenges that might be related to culture, for instance in terms of cross-gender communications, geographical issues (distance), time (cancellation), participants' concerns about privacy, hesitation from some participants and issues of compensation.

It was difficult to have face-to-face interviews across the genders. Therefore, male participants were asked to choose if they wanted to conduct interviews by telephone or through online chat. The workshops comprised only female participants as it was not possible for them to be mixed gender or to ask someone else to conduct sessions for male participants.

In terms of distance and geographic challenges, conducting interviews in different locations in Saudi Arabia could be time consuming and costly, requiring travel between locations sometimes lasting more than 1 hour. Also, sometimes the interviewee would be late or not arrive at all. Therefore, online chat interviews were more convenient. For the workshops, the potential for non-attendance, which did happen, was taken into account by inviting larger numbers than required.

Consistent with research guidelines (e.g. Code of Good Practice in Research and Handling of Personal and Sensitive Data guidelines from Newcastle university), the participants were informed that their information and details would be totally anonymised. Also, they were told that in this study they were the experts in terms of how to use the platforms, helping give them the confidence to explain even the simplest details or activities. In some cases, the participants asked if they had provided the "right" answer and therefore they were informed before the interviews that there were no correct or wrong answers or correct or wrong behaviours and that the answers only aimed to understand the use of s-commerce.

For the design critique workshops, the participants were informed that the design had been worked on by previous participants and they were free to critique it. This helped the participants feel free to critique the design.

Finally, as already noted, many participants rejected monetary compensation in the form of a gift card. They kept saying that they wanted to participate on a voluntary basis. In some cases, I therefore gave them chocolate, which was more acceptable.

Transcription and analysis

Transcribing interviews and workshops takes time and there is no readily available software that can do it effectively for Arabic. Some websites claim to provide auto-transcription in Arabic, but it does not match at all. Therefore, transcription was

undertaken manually, which was very time-consuming: for each hour of recording, it might take three to four hours to transcribe, especially with background noise. Interviews undertaken in online chat were effectively already transcribed, being in text, which meant less effort. However, the interactions in the workshops were more difficult to transcribe as there were multiple people and sometimes two groups were talking at the same time. To make it possible to distinguish what the two groups were saying, individual recording devices were used for each.

Conducting online chat interviews would potentially save time on transcription as the interviews would already be in text form (unless using voice notes). In contrast, the telephone and face-to-face interviews needed to be transcribed. The telephone interviews were more difficult to transcribe due to background noise, both on the participants' end and speaker interference. The face-to-face interviews were easier to transcribe as we were both (interviewer and interviewee) in the same location and there was minimal or even no background noise, especially when the interviews were conducted in meeting rooms or offices.

The language used in this study was Arabic and ATLAS.ti was used to help analyse the data as the program supports Arabic. In contrast, NVivo, often used for qualitative analysis, was not suitable for use with Arabic script and several issues arose when trying to employ it, such as the imported file not showing all the text and it skipping when trying to code a phrase.

Qualitative coding and word count

Word counts and codes were tested to check if there were significant differences between different data collection methods and forums. A global word count was used without deleting repetition. Four comparisons were made as follows: (i) online chat and telephone; (ii) online chat and face to face; (iii) face to face, online chat and telephone; (iv) co-design workshops and design critique workshops.

Comparison of online chat and telephone interviews (Study 1)

As can be seen from Table II, although the word count for online chat is greater, the number of codes for telephone is higher. While the difference in the word count is significant ($p = 0.001$), the difference in the number of codes is not ($p = 0.101$) and therefore, there are no significant differences in the information obtained from the two methods.

Interview form	Word count	Codes
Online chat	1426	43
Telephone	741	49

Table II: Median scores for online chat and telephone interview data

Comparison of online chat and face-to-face interviews (Study 3)

As is apparent in Table III, the word count and number of codes were both higher in the online chat interviews. The differences in both instances are significant ($p = 0.001$). This result suggests that online chat can obtain more information than face-to-face interviews. It should be noted that in this case, the interviews entailed asking and answering questions and there were no additional activities.

Interview form	Word count	Codes
Online chat	3297	111
Face to face	2146	85

Table III: Median scores for online chat and face-to-face interviews

Comparison of online chat, telephone and face-to-face interviews including activities (Studies 1 and 2)

Table IV shows that the word count for the face-to-face interviews (including activities) and the number of codes are higher than for the telephone and online chat interviews. Moreover, the difference in word count between face-to-face interviews and online chat is significant ($p = 0.001$) and the difference in the number of codes is between face-to-face interviews and telephone interviews is also significant ($p = 0.005$). Thus, it appears that face-to-face interviews including activities can provide richer data and greater detail than other methods of interviewing.

Interview form	Word count	Codes
Online chat	1426	43
Telephone	741	49
Face to face	1868	60

Table IV: Median scores for online chat, telephone and face-to-face interviews with activities

Comparison of co-design and critique design workshops

As shown in Table V, the co-design workshops present both a higher word count and a greater number of codes than the critique design workshops, a difference that is significant ($p = 0.001$). This is most likely because the co-design workshops included more discussion of previous experiences and stories of shopping issues that participants might face. In the critique design workshops, the participants were focused on evaluating the design in terms of what they liked/disliked and suggestions.

Interview form	Word count	Codes
Co-design workshop	5250	290
Critique design workshop	3681	149

Table V: Median scores for co-design workshops and critique design workshops

Discussion and conclusion

Semi-structured interviews including no additional activities can be undertaken through online chat, by telephone or face to face. Based on the results, online chat and telephone are better at eliciting information than face-to-face interviews. However, if semi-structured interviews include activities, participants can be encouraged to provide more information.

The word count is not particularly important as sometimes the participants repeat what they have said or amplify rather than just giving direct answers. In contrast, the number of codes potentially reflects the provision of greater information.

Online chat interviews are the cheapest option and require less effort on the part of the researcher as they can be undertaken anytime and anywhere, there are no issues with cross-gender communication, the participants and researcher are not stressed and participants are comfortable about saying whatever they want. It is also beneficial to use online chat as participants can share photographs, providing rich data and additional details.

Telephone interviews can be costly in terms of time and require total focus. Not only is it necessary to address potential noise and signal problems, but also participants may be busy or distracted and cut questions, which is distracting. Recordings can be unclear and they can be difficult to transcribe. On the other hand, they can provide rich data.

Face-to-face interviews not including activities are the least recommended means of collecting data as they require considerable effort. They can be costly because of having to travel to the meeting place, waiting for participants to show up and transcribing the interviews. The data gathered could be less informative than those obtained through telephone interviews and online chat. However, if the interviews include activities, such as sharing maps, using mobile phones to recall shopping experiences, or using picture cards to discuss the features, face-to-face interviews can provide richer data and more details than telephone interviews or online chat. It is important to bear in mind the cost, time and effort in selecting the optimal method.

Finally, co-design and critique design workshops can be beneficial in informing on the optimal design. The co-design workshops included a wealth of information as the participants explained their previous experiences. However, in the critique design workshops, the participants were focused more on the existing design and commenting on aspects related solely to the design. It is necessary to

employ a mix of methods to learn about users' needs and experiences. Also, it is useful to incorporate more than one method to validate previous results (Steen et al., 2011). Table VI summarises the advantages and disadvantages of the methods implemented in this research.

Table VI: Summary of advantages and disadvantages of methods conducted in this research.

Method	Advantages	Disadvantages
Semi-structured interviews - by text	<ul style="list-style-type: none"> • Preferred by participants as they can take place while they are at home with their family. • There is no need to travel • Can be done over wide geographic distances • More comfortable for cross-gender communication • Can keep going without the feeling that they want to end it • No need to transcribe 	<ul style="list-style-type: none"> • Can be slower than other methods • May take a longer time to finish the interview
Semi-structured interviews - by voice	<ul style="list-style-type: none"> • More comfortable for cross-gender communication • Preferred by participants as they do not need to travel and meet in a specific place • Do not need a location • Saves researcher's time (travelling to the location and waiting for the participants) 	<ul style="list-style-type: none"> • Could be unclear (e.g. signalling issues) • Could have background noise (e.g. children, family members, car, people) • Might want to end the call sooner than desirable • Hard to transcribe
Semi-structured interviews - face to face	<ul style="list-style-type: none"> • Interviewee focus is mostly on the interview (depending on the location) 	<ul style="list-style-type: none"> • Needs a location • Takes time to travel to the location and wait for the participants • Takes time to transcribe
Online observation	<ul style="list-style-type: none"> • Helps to provide overall information about the general interaction and activities • Helps to support the data collected by other methods such as interviews 	<ul style="list-style-type: none"> • If the platform does not make it possible to collect data through a tool, it is more difficult to conduct observation • It is not possible to observe private interactions and activities • It can be time-consuming as the observation needs to take place over a period of time (several hours at a time over several weeks/months)
Semi-structured interviews - face to face with activities (sharing map)	<ul style="list-style-type: none"> • Helps participants provide more details • Helps participants remember the details of previous experiences • Participants are more open to talking as they look at the drawing and remember to mention who they share experiences with • Can help make interviews interesting and keep the participant engaged 	<ul style="list-style-type: none"> • Time-consuming to travel to the location and to transcribe • Need for a location • Difficult to conduct cross-gender interviews
Diary	<ul style="list-style-type: none"> • Provides additional information (experiences) to supplement interviews 	<ul style="list-style-type: none"> • Participants forget to share diary entries and the researcher needs to send reminders

Continuation of Table VI

Method	Advantages	Disadvantages
Co-design workshops	<ul style="list-style-type: none"> • Helps to understand needs and issues • Having more than one participant can bring new ideas and insights to the discussion • Discussing previous experiences and challenges • Using card sorting helps participants discuss issues in greater depth and show what they want and need 	<ul style="list-style-type: none"> • Some participants might not wish to share thoughts, so the researcher should try to keep the discussion going and include all the participants • Some participants might start to talk about things not related to the workshop; then the researcher should again bring the discussion back on topic • Participants might not want to sketch or draw
Critique workshops	<ul style="list-style-type: none"> • Participants will focus on critiquing the design and specify what they like • Can help to improve the design • Helps to check how they interact with the system and they might realise the need for features not previously considered before interacting with the platform 	<ul style="list-style-type: none"> • Participants might start to focus on specific details, such as colours or photos • Some participants might start to talk about things not related to the workshop; then the researcher should again bring the discussion back on topic • Some participants might not share their thoughts, so the researcher should try to keep the discussion going and include all the participants
SUS	<ul style="list-style-type: none"> • Helps identify if there is general acceptance from the users • Identifies the extent to which the system is easy to use and understandable 	<ul style="list-style-type: none"> • Might not be accurate in testing an interactive prototype as not all the functions can be used • Not a useful method to be used alone as the researcher needs to know the difficulties that the participants face or issues with the system

References

- Abed, S. S., Y. K. Dwivedi, and M. D. Williams (2015): 'Consumers' perceptions of social commerce adoption in Saudi Arabia'. In: *Conference on e-Business, e-Services and e-Society*. pp. 133–143.
- Abokhodair, N. and S. Vieweg (2016): 'Privacy & social media in the context of the Arab Gulf'. In: *Proceedings of the 2016 ACM conference on designing interactive systems*. pp. 672–683.
- Al-Dawood, A., N. Abokhodair, H. El Mimouni, and S. Yarosh (2017): "' Against Marrying a Stranger" Marital Matchmaking Technologies in Saudi Arabia'. In: *Proceedings of the 2017 Conference on Designing Interactive Systems*. pp. 1013–1024.
- Al-Dawood, A., S. Alhajhusein, and S. Yarosh (2021): 'Saudi Arabian Parents' Perception of Online Marital Matchmaking Technologies'. *Proceedings of the ACM on Human-Computer Interaction*, vol. 4, no. CSCW3, pp. 1–32.
- Al-Saggaf, Y. (2011): 'Saudi females on Facebook: An ethnographic study'. *International Journal of Emerging Technologies and Society*, vol. 9, no. 1, pp. 1.
- Al-Saggaf, Y., K. Williamson, et al. (2004): 'Online communities in Saudi Arabia: Evaluating the impact on culture through online semi-structured interviews'. In: *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*.

- Al-Tit, A. A., A. Omri, and T. B. Hadj (2020): 'The driving factors of the social commerce intention of Saudi Arabia's online communities'. *International Journal of Engineering Business Management*, vol. 12, pp. 1847979019899746.
- AlArfaj, A. A. and E. Solaiman (2019): 'Investigating commercial capabilities and trust in social media applications for entrepreneurs'. In: *Proceedings of the 9th International Conference on Communities & Technologies-Transforming Communities*. pp. 65–75.
- AlArfaj, A. A. and E. Solaiman (2021): 'Consumer Needs and Design Practices for Trusted Social Commerce Platforms'. In: *IFIP Conference on Human-Computer Interaction*. pp. 205–226.
- AlArfaj, A. A., E. Solaiman, and L. Marshall (2019): "'Why would you buy from a stranger?'" Understanding Saudi citizens' Motivations and Challenges in Social Commerce'. In: *IFIP Conference on Human-Computer Interaction*. pp. 711–732.
- Alghamdi, D., I. Flechais, and M. Jirotko (2015): 'Security practices for households bank customers in the Kingdom of Saudi Arabia'. In: *Eleventh Symposium On Usable Privacy and Security ({SOUPS} 2015)*. pp. 297–308.
- AlGhamdi, E. and D. Reilly (2013): 'Social media based business in Saudi Arabia'. In: *Dalhousie University*. pp. 1–3.
- Almakrami, H. A. A. (2015): 'Online self-disclosure across cultures: A study of Facebook use in Saudi Arabia and Australia'. Ph.D. thesis, Queensland University of Technology.
- Alotaibi, T. S., A. A. Alkhathlan, and S. S. Alzeer (2019): 'Instagram shopping in Saudi Arabia: what influences consumer trust and purchase decisions'. *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 11.
- Alsheikh, T., J. A. Rode, and S. E. Lindley (2011): '(Whose) Value-Sensitive design: a study of long-distance relationships in an Arabic cultural context'. In: *Proceedings of the ACM 2011 conference on Computer supported cooperative work*. pp. 75–84.
- Anari, S. M. M., R. Ibrahim, and A. Bakri (2014): 'Understanding factors on the customer intention behavior through Facebook commerce: a conceptual model'. In: *2014 International Symposium on Technology Management and Emerging Technologies*. pp. 107–112.
- Baethge, C., J. Klier, and M. Klier (2016): 'Social commerce—state-of-the-art and future research directions'. *Electronic Markets*, vol. 26, no. 3, pp. 269–290.
- Baxter, K., C. Courage, and K. Caine (2015): *Understanding your users: a practical guide to user research methods*. Morgan Kaufmann.
- Bowers, J. M. (1991): 'The Janus Faces of Design: Some Critical Questions for CSCW'. In: J. M. Bowers and S. D. Benford (eds.): *Studies in Computer Supported Cooperative Work: Theory, Practice and Design*. Amsterdam, etc., pp. 333–350, North-Holland.
- Busalim, A. H. et al. (2016): 'Understanding social commerce: A systematic literature review and directions for further research'. *International Journal of Information Management*, vol. 36, no. 6, pp. 1075–1088.
- Chen, L. (2010): 'Understanding buyers' social information needs during purchase decision process'. In: *Proceedings of the 12th International Conference on Electronic Commerce: Roadmap for the Future of Electronic Business*. pp. 15–22.

- Dillahunt, T. R. and A. R. Malone (2015): 'The promise of the sharing economy among disadvantaged communities'. In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. pp. 2285–2294.
- Dimond, J. P., C. Fiesler, B. DiSalvo, J. Pelc, and A. S. Bruckman (2012): 'Qualitative data collection technologies: A comparison of instant messaging, email, and phone'. In: *Proceedings of the 17th ACM international conference on Supporting group work*. pp. 277–280.
- Dye, M., A. Antón, and A. S. Bruckman (2016): 'Early adopters of the Internet and social media in Cuba'. In: *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*. pp. 1295–1309.
- Esmaeili, L. and S. A. Hashemi G (2019): 'A systematic review on social commerce'. *Journal of Strategic Marketing*, vol. 27, no. 4, pp. 317–355.
- Evans, H. I., M. Wong-Villacres, D. Castro, E. Gilbert, R. I. Arriaga, M. Dye, and A. Bruckman (2018): 'Facebook in Venezuela: understanding solidarity economies in low-trust environments'. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. pp. 1–12.
- Flechais, I., M. Jirotko, and D. Alghamdi (2013): 'In the balance in Saudi Arabia: security, privacy and trust'. In: *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. Acm, pp. 823–828.
- Friedrich, T., S. Overhage, and S. Schlauderer (2016): 'The more the better? Exploring the relationship between social commerce feature intensity, social factors, and consumers' buying behavior'. In: *Thirty Seventh International Conference on Information Systems, Dublin*.
- Gerson, E. M. and S. L. Star (1986): 'Analyzing due process in the workplace'. *ACM Transactions on Office Information Systems*, vol. 4, no. 3, pp. 257–270.
- Gheitasy, A., J. Abdelnour-Nocera, and B. Nardi (2015): 'Socio-technical gaps in online collaborative consumption (OCC) an example of the etsy community'. In: *Proceedings of the 33rd Annual International Conference on the Design of Communication*. pp. 1–9.
- Greenberg, S. and B. Buxton (2008): 'Usability evaluation considered harmful (some of the time)'. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. pp. 111–120.
- Han, H., H. Xu, and H. Chen (2018): 'Social commerce: A systematic review and data synthesis'. *Electronic Commerce Research and Applications*, vol. 30, pp. 38–50.
- Hillman, S., A. Forghani, C. Pang, C. Neustaeder, and T. K. Judge (2014): 'Conducting interviews with remote participants'. In: *Studying and Designing Technology for Domestic Life*. pp. 11–53, Morgan Kaufmann.
- Hillman, S., C. Neustaeder, C. Pang, and E. Oduor (2013): '" Shared joy is double joy" the social practices of user networks within group shopping sites'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 2417–2426.
- Huang, Z. and M. Benyoucef (2013): 'From e-commerce to social commerce: A close look at design features'. *Electronic Commerce Research and Applications*, vol. 12, no. 4, pp. 246–259.
- Huang, Z., S. Y. Yoon, and M. Benyoucef (2012): 'Adding social features to e-commerce'. In: *Proceedings of the Conference on Information Systems Applied Research ISSN*, Vol. 2167. p. 1508.

- Ji, C., Q. Chen, and N. Zhuo (2020): 'Enhancing consumer trust in short food supply chains: The case evidence from three agricultural e-commerce companies in China'. In: *Journal of Agribusiness in Developing and Emerging Economies*. Emerald Publishing Limited.
- Johansen, R. (1988): *Groupware. Computer Support for Business Teams*. New York and London: The Free Press.
- Krimly, R. (2019): 'Revamped guardianship laws usher in a new era for Saudi women'. <https://english.alarabiya.net/en/features/2019/08/03/Revamped-guardianship-laws-usher-in-a-new-era-for-Saudi-women>. Accessed: 2021-09-30.
- Lazar, J., J. H. Feng, and H. Hochheiser (2017): *Research methods in human-computer interaction*. Morgan Kaufmann.
- Maeng, W., J. Yoon, H. Ahn, and J. Lee (2016): 'Can mobile instant messaging be a useful interviewing tool? A comparative analysis of phone use, instant messaging, and mobile instant messaging'. In: *Proceedings of HCI Korea*. ACM, pp. 45–49.
- Makki, E. and L.-C. Chang (2015): 'Understanding the effects of social media and mobile usage on e-commerce: an exploratory study in Saudi Arabia'. *International management review*, vol. 11, no. 2, pp. 98–109.
- Martin, B., B. Hanington, and B. M. Hanington (2012): *Universal methods of design: 100 ways to research complex problems*. 100 Cummings Center, Suite 406L, Beverly, MA 01915: Rockport Publishers.
- Moser, C., P. Resnick, and S. Schoenebeck (2017): 'Community commerce: Facilitating trust in mom-to-mom sale groups on Facebook'. In: *Proceedings of the 2017 CHI conference on human factors in computing systems*. pp. 4344–4357.
- Muller, M. J. and A. Druin (2012): 'Participatory design: the third space in human-computer interaction'. In: *The Human-Computer Interaction Handbook*. CRC Press, pp. 1125–1153.
- Nassir, S. and T. W. Leong (2017): 'Traversing boundaries: Understanding the experiences of ageing Saudis'. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. pp. 6386–6397.
- Nassir, S. and T. W. Leong (2018): 'Conducting Qualitative Fieldwork with Ageing Saudis: A Visual Diary'. In: *Proceedings of the 2018 Designing Interactive Systems Conference*. pp. 427–439.
- Nylander, S. and Å. Rudström (2011): 'Questions, inspiration, feedback, and contributions: How entrepreneurs network online'. In: *Proceedings of the 5th International Conference on Communities and Technologies*. pp. 128–137.
- Rumpf, H. (1990): 'The characteristics of systems and their changes of state disperse'. In: *Particle Technology*, Chapman and Hall. pp. 8–54, Springer: Berlin/Heidelberg, Germany.
- Sanders, E. B.-N. and P. J. Stappers (2008): 'Co-creation and the new landscapes of design'. *Co-design*, vol. 4, no. 1, pp. 5–18.
- Sheikh, Z., T. Islam, S. Rana, Z. Hameed, and U. Saeed (2017): 'Acceptance of social commerce framework in Saudi Arabia'. *Telematics and Informatics*, vol. 34, no. 8, pp. 1693–1708.
- Shin, D.-H. (2013): 'User experience in social commerce: in friends we trust'. *Behaviour & information technology*, vol. 32, no. 1, pp. 52–67.

- Steen, M., M. Manschot, and N. De Koning (2011): 'Benefits of co-design in service design projects'. *International Journal of Design*, vol. 5, no. 2.
- Suraworachet, W., S. Preamsiri, and N. Cooharajanone (2012): 'The study on the effect of facebook's social network features toward intention to buy on f-commerce in Thailand'. In: *2012 IEEE/IPSJ 12th International Symposium on Applications and the Internet*. pp. 245–250.
- Tomitsch, M., C. Wrigley, M. Borthwick, N. Ahmadpour, J. Frawley, A. B. Kocaballi, C. Núñez-Pacheco, and K. Straker (2018): *Design. think. make. break. repeat. A handbook of methods*. Bis Publishers.
- Voida, A., E. D. Mynatt, T. Erickson, and W. A. Kellogg (2004): 'Interviewing over instant messaging'. In: *CHI'04 extended abstracts on Human factors in computing systems*. pp. 1344–1347.
- Zhang, K. Z. and M. Benyoucef (2016): 'Consumer behavior in social commerce: A literature review'. *Decision support systems*, vol. 86, pp. 95–108.

Max Krüger¹, Marc Gerbracht¹, Nico Vitt¹, Muhamed Kudic¹, Michael Ahmadi¹, Alexander Boden^{2,3}, Felicitas Offergeld², Martin Stein², Christoph Kotthaus⁴, David Unbehaun¹, Volker Wulf¹ (2022): Travelling Artefacts: Lessons Learned from Interventions in a Regional Innovation Ecosystem. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_ep06

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_ep06

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

Travelling Artefacts: Lessons Learned from Interventions in a Regional Innovation Ecosystem

Max Krüger¹, Marc Gerbracht¹, Nico Vitt¹, Muhamed Kudic¹, Michael Ahmadi¹, Alexander Boden^{2,3}, Felicitas Offergeld², Martin Stein², Christoph Kotthaus⁴, David Unbehaun¹, Volker Wulf¹

¹Institute of Information Systems and New Media, Universität Siegen, Germany;

²Fraunhofer-Institute for Applied Information Technology FIT, Schloss Birlinghoven Sankt Augustin, Germany; ³Institute for Consumer Informatics, Bonn-Rhein-Sieg University of Applied Science H-BRS, Sankt Augustin, Germany; ⁴Aventum GmbH, Germany

maximilian.krueger, marc.gerbracht, nico.vitt, muhamed.kudic, michael.ahmadi, david.unhaun, volker.wulf@uni-siegen.de; Alexander.boden@h-brs.de; felicitas.offergeld, martin.stein@fit.fraunhofer.de, christoph.kotthaus@aventum.de

Abstract. Regions and their innovation ecosystems have increasingly become of interest to CSCW research as the context in which work, research and design takes place. Our study adds to this growing discourse, by providing preliminary data and reflections from an ongoing attempt to intervene and support a regional innovation ecosystem. We report on the benefits and shortcomings of a practice-oriented approach in such regional projects and highlight the importance of relations and the notion of spillover. Lastly, we discuss methodological and pragmatic hurdles that CSCW research needs to overcome in order to support regional innovation ecosystems successfully.

Introduction

In the last decade interest in the role of projects within regional political, cultural and economic developments within the CSCW and wider HCI community has grown. This includes for example research into the notion of innovation in different geographic contexts and its relation to economic development and national identity (Avle and Lindtner (2016)) or exploration of the collaborative creation of visions for manufacturing in Taiwan (Freeman et al. (2018)). The importance of rural areas has also increasingly come into focus of CSCW research (Hardy et al. (2019b,a)), exploring the specificities of HCI in rural areas in a series of workshops (Hardy et al. (2018)), papers (e.g. (de Castro Leal et al. (2021))) and Special Issues of journals (Makoto Su et al. (2021)). We build on this literature by contributing an investigation into regional learning processes about digitization issues with small and medium sized enterprises in a rural-industrialized area in Germany. While previous studies have consisted predominantly of empirical investigations into the role of regional ecosystems for CSCW and HCI, this paper describes an attempt to intervene and shape such ecosystems through practice- and action-oriented research and design projects, in a rural yet industrialized area in Germany.

Germany is characterized by a considerable number of rural regions, which at the same time exhibit a high level of industrialization. They provide skilled labor, the necessary space and affordable real estate, making them home to the majority of Small and Medium-sized Enterprises (SMEs) which characterize German's industrial sector. Thus, they constitute an important part in the country's economy. This leads to the fact that some of the country's areas are rural yet also highly industrialized. While these regions are typically still prosperous, they are confronted with increasing pressure to modernize their traditionally energy-intensive industries.

Despite the importance of rural regions for the local economy and the fact that "equality of quality of life" between urban and rural areas is included in German constitution (Art. 72(2), GG), they are facing significant challenges. The population of many rural areas is shrinking: birth rates are decreasing, youth is moving to urban areas, rural areas do not manage to attract and welcome people from outside or create sufficient access to the labor market (Swiaczny (2015); Vogelgesang et al. (2018)). Amongst others, responsible factors are a lack of educational opportunities, restricted employment opportunities, especially for university graduates and a general cultural closedness (Bundesministerium für Ernährung und Landwirtschaft (BMEL) (2015)). A variety of options might be available to counter the potential negative consequences for rural regions. One option that the German government actively promotes is to support SMEs in their efforts to digitize their organization, process and products through a network of so-called 'Centers of Competency'. These centers offer a variety of support mechanisms to companies, from low-threshold informational events to trainings, consisting of several in-depth workshops, to hands-on projects which introduce

and appropriate digital technologies in a prototypical manner into the company context. This paper reports on the experiences and insights gained from work in one of these centers.

To succeed with our initiative, we anticipated to require a ‘broad’ stakeholder and context-sensitive view, to receive a more holistic picture (Ogonowski et al. (2013)) not only of the individual companies we would interact with but of the regional innovation ecosystem. This seems important for several reasons in our context: While the different activities and services the center offers are usually directed at specific companies, the regional innovation ecosystem they operate in consists of a broad variety of actors. Especially in rural regions, such actors always have close and long-standing relationships. Important roles are filled by various unions and their representatives, employer associations such as chambers of commerce and craft, government organizations and other public administrative institutions that offer resources and support to companies. Any project attempting to address and positively shape an ecosystem would need to take this constellation of actors and the relationships between them into account. This is why, the center under investigation employs a practice-based and open co-design-oriented approach to achieve this (Følstad et al. (2009)), where relevant stakeholders are involved in innovation activities (Chesbrough et al. (2006)) and that allows long-term cooperation (Wulf et al. (2015)). Such a ‘practice-based approach’ aims at fulfilling societal needs, triggering social change by addressing ‘real world problems’ (Wulf et al. (2018)). The paper presents early results from this intervention and reflects on however such an approach is appropriate for the attempted goal to intervene in a region, instead of single company, department or context, and the perceived benefits and challenges of it.

To describe our activities within the field, we adopted a meta-research or ‘research on research’ approach (Dachtera et al. (2014); Randall et al. (2018)), building in the internal evaluation activities of the Center. Our contribution to the CSCW discourse is two-fold: Firstly, we expand on the current discourse on regional innovation in HCI by reporting on an attempt to not only analyze but intervene in and shape a regional innovation ecosystem through practice-oriented CSCW activities such as co-design workshops, which we describe below in more detail. This, we believe presents a potentially fruitful and new line of investigation and intervention for HCI research, with effects beyond single organizational contexts. Secondly, by building explicitly on a practice-oriented approach to HCI, we outline the benefits and challenges of such an approach when focusing on a region instead of an individual practice-context. We find here that ‘spillover’ is a crucial element of regional activities, as the effects of CSCW interventions ‘travel’ between departments, companies and even sectors and are not bound to the specific practice context where they originated. The question for CSCW projects that wish to address regional contexts then becomes how such spillover effects can be supported.

In the following, we will briefly discuss the existing literature on regional development and innovation, with a specific focus on the study of such constellations in HCI.

Related Work - Innovation Ecosystems & HCI

Several studies have investigated the functions of regional innovation ecosystems from various disciplinary perspectives. An influential study was published by Saxenian, who investigated the innovation ecosystem of Silicon Valley, highlighting the very specific role of universities such as Stanford and Berkeley in creating a workforce highly skilled for local tech companies and in transforming the valley from farmland to a high-tech region (Saxenian (1994); Adams (2005); Tajnai (1996)). The dynamics of innovation processes in geographical regions has for a long time been the interest and focus the economic sciences. Here, interdisciplinary innovation network research has been a vibrant field of research (Ozman (2009); Bergenholtz and Waldstrøm (2011); Kudic et al. (2021)). This literature observes a broad variety of motives to collaborate in with other network partners in regional innovation ecosystems, such as cost savings and risk reduction (Hagedoorn (1993)), time savings (Mowery et al. (1996)), access to national and international markets (Hagedoorn (1993)), status and reputation building (Gulati (1998)), knowledge access (Grant and Baden-Fuller (2004)) and interorganizational learning (Hamel (2011)). Within this broader discourse the concept of the Regional Innovation System is perhaps most relevant to the goals of the paper. RIS can be differentiated according to Asheim (1998) and Pyka et al. (2019) into three types: (i) territorially embedded regional innovation systems, (ii) regionally networked innovation systems, and (iii) regionalized national innovation systems (Asheim and Coenen (2005)). The various approaches differ in the kind of relations they describe and distinguish innovation ecosystems according to these relations.

In recent years, international research within the wider HCI community has increasingly focused on the workings of regional innovation networks in various countries and the role of HCI projects within them. Several studies have for example highlighted a growing importance of “maker spaces” and other “innovation hubs” in Shenzhen, China (Lindtner et al. (2014)) or Taiwan (Lindtner et al. (2016); Freeman et al. (2018)). Such spaces play a role of increasing importance in the support of technological (and social) innovation outside of traditional research and development labs of firms or universities (Lindtner et al. (2014)). Similar developments have been studied in a variety of countries (Avle and Lindtner (2016); Csikszentmihalyi et al. (2018)). Through these studies it becomes apparent that attempts to transfer innovation models and methods such as Hackathons, “Design Thinking” or Start-Up Accelerators, that predominantly stem from California’s Silicon Valley and are portrayed as universally successful across cultures, contexts and countries to places as diverse as Jamaica, Ghana or China (Avle et al. (2017)) or India (Irani (2019)), are highly problematic. As they do not

take into account local conditions and practices, they are either unsuccessful in the support of innovation or even hinder innovation as they impede the work of local innovators instead of alleviating it (Irani (2019); Irani et al. (2010)). Accordingly, these studies provide an important contribution to the HCI community in analyzing the wider ecosystems in which our work takes hold and is a crucial backdrop to our work. Several studies provide insights on what it might take, methodologically, to intervene in innovation ecosystems. This includes work by Dachtera et al. (Dachtera et al. (2014)), who discussed the opportunities and challenges of university-industry partnerships. They point out that such collaborations and their increasing frequency led to a new paradigm of knowledge production, labelled post-academic (Ziman (2000)) or Mode 2 (Nowotny et al. (2013)) knowledge production. They draw out several challenges to such endeavors, including organizational and epistemological differences between the partners, that need to be taken into account when targeting an innovation ecosystem, including the companies it entails. Some projects have explicitly focussed on the development of digital infrastructures that support collaboration between organisations within a region, such as shared mobility solutions (see e.g. Stein (2017)). Fischer et al. (2007) have further investigated the role of knowledge creation in collaborations between universities and regional innovation networks. Such partnerships, they argue, have the potential to be intensely beneficial to students, providing opportunities to be lifelong learners, as well as for regional innovation, if universities accomplish to take "the importance of industrial practice and social networks into account". Our study adds to this growing discourse, by providing preliminary data and reflections from an ongoing attempt to intervene, shape and support a regional innovation ecosystem.

Background & Context

The Region

Our activities took place in one of the oldest industrial areas of central Europe, characterized by small and medium-sized companies (250 employees or less), the so-called "Mittelstand", with the exception of a few larger companies (several thousand employees and production sites in different countries). About half of the gross economic value created in this region comes out of the manufacturing sector, including areas such as automotive suppliers, machine tools, plant production and engineering, plastics processing, home as well as electrical equipment such as batteries, lights, wire and transformers. The regional economy is strongly oriented towards export, with about 160 companies in the region being so-called "hidden champions" in global niche markets.

Due to the climate crisis, a changing global economic landscape and the demands to digitize their products, processes, and services, the regional industry is in the middle of a transformation process. This comes with potentially far reaching social and economic consequences for the inhabitants. Since these traditional

industries are energy-intensive, the region's companies are under pressure to rearrange their production and logistical procedures to reduce CO2 emissions – a change in which digitization can play a crucial role (Strüker et al. (2021)). However, the regional SMEs often lack the resources to make necessary and substantial investments in digitization. This runs the risk of dire economic consequences for local companies, which in turn could also negatively affect the region as a whole.

To overcome digitization hurdles which SMEs face nationwide, the Federal Ministry of Energy and Economy has financed 26 so-called “Centers of Competency” across the entire country, with either a regional or thematic focus on specific aspects of digitization. These centers are tasked to inform SMEs on digitization and actively support them in taking first steps towards digitization within their organizations by offering practical and context-specific support. Even though many of these centers are based at and run by universities, they are explicitly not *research* but "*transfer*"-oriented projects. Their goal is to support companies by 'transferring' knowledge to companies and to enable them to *apply* digital technologies, which means that activities need to be highly practical and deliver concrete results to the participating companies. In this paper we report on the activities of one of these centers.

The Centre of Competency

The Center of Competency under investigation in this paper consists of a large consortium, including several educational institutions and universities of the region and various research groups of the region's central university (located in the largest city of the region), providing expertise in Human Computer Interaction, Economics as well as various sub-disciplines of mechanical engineering. Each of the consortium partners conducts their own activities, focusing on adjacent yet different (sub-)regions, but naturally cooperation frequently occurs, for example when members of one institution run a workshop as part of a series organized primarily by another consortium partner. The focus of this paper, however, lies primarily on the activities of the city's university, acknowledging that it is embedded in the larger infrastructure of the center.

The main goal of the specific Competency Centre is to decrease the “digital gap” of regional SMEs compared to large companies, in line with the requirements of the region and the funding objectives of the project. Despite that, the Center occasionally also works with the larger companies of several thousand employees. The publicly funded initiative began its work in late 2017 and since then has been implementing measures towards the creative, economically sustainable and technologically innovative digitization of regional SMEs. Crucially, in all its activities the center aims to shape a form of digitization that embodies the spirit of “Industrial Relations” (Haipeter (2012)) and is therefore largely employee-centered. Furthermore, and especially important for this paper, the activities of the Center are grounded in the assumption, that digitization projects

should support qualified human work, rather than attempt to standardize or automate it (Ludwig et al. (2016); Wulf et al. (2018)), which is largely in line with the wider interest of CSCW into (work) practices ((Lanamäki and Väyrynen (2016))).

In its core, the Centre focuses on the implementation and application of digital technologies in companies, foregrounding Human Computer Interaction. Its interest lies not in the technologies themselves, but in their application and appropriation, and the accompanying changes in the work practices within the organizations and the wider innovation ecosystem. This means that the Center is characterized by the conviction, that workers need to play a central role in digitization measures and should be included in all steps, as their knowledge and practices are crucial both to the performance of the company as well as to the design and appropriation of technological artefacts. Specifically, the creativity, existing expertise, knowledge and competencies of employees, that in this specific region have often been working in a company for many years, are considered crucial to the companies' success. They should not be replaced but augmented by digital technologies, thereby supporting a digitization in the tradition of Industrial Relations (Ludwig et al. (2016); Haipeter (2012)).

Epistemological and Methodological Foundations of the Center

As indicated in the introduction, underlying all activities is an epistemological and methodological commitment to the central role of practices, following the practice-based paradigm of CSCW and HCI research (Schmidt and Bannon (2013); Bjørn et al. (2016); Kuutti and Bannon (2014)). Following such a practice-oriented paradigm, technological artefacts are not simply used, but appropriated by users to their specific context (Pipek and Wulf (2009)). Technologies thus become embedded into human action and practices, a process which is highly dynamic, nuanced and contextual (Ackerman (2000)). Technological artefacts are necessarily abstractions, as they are created based on assumptions of designers and developers and are shaped according to the situative requirements known to them. In practice then, such artefacts are appropriated by their users, and thereby re-contextualised (Rohde et al. (2017); Pipek and Wulf (2009); Stevens et al. (2009, 2010)). 'Use' is therefore an active and creative process, as artefacts are given meaning in their specific context of application by those that apply and appropriate them (Suchman (1993, 2002)). If the use-context, as understood and imagined by the designers, diverges too much from the actual use-context of the users, these will face considerable difficulties to integrate the new tools into their work practice. To minimize this discrepancy, one way to conduct a practice-based approach to design is to structure research into "design case studies", each of which includes an iterative combination of empirical investigation into the specific practices at hand, a participatory co-design-intervention in the form of a technological artefact, and investigating and supporting processes of appropriation of the created artefact to the specific

context (Wulf et al. (2011, 2015); Stevens et al.). This iterative and participatory process aims to overcome the asymmetry of knowledge (Fischer (2000); Rittel (1984)) that exists between the different actors, such as academic researchers and practitioners, and create shared knowledge about the practice context at hand and the design possibilities, a symmetry of knowledge (Fowles (2000)). The aim of the Center's activities is an intervention in such practices, mostly via the deployment of socio-technical artefacts, although, as we shall make clear, interventions often also take other forms, and a description of practices is not the central concern of this paper. Underlying all the interventions is not a commitment to the format of a design case study, but to the understanding of technologies as socio-technical artefacts, given meaning in use, and to the requirement that members of the contexts we intervene in participate in the digitization process.

Knowledge exchange is furthermore central to the works of the center, as it is tasked to increase SME's competency to address and implement digitization measures successfully. Traditionally learning is associated with the idea of transferring knowledge from an expert to a learner, and this is perhaps the root of the idea of a 'transfer project,' which is at least common term for the kind of project we are discussing here. We divert from this idea, building on socio-cultural theories of learning (see e.g. (Lave and Wenger (1991))) following the example of Fischer et al. (Fischer et al. (2007)). Learning in the center of competency is understood as a process of co-creating knowledge between different actors, without pre-defined and static roles of expert and learner. In this understanding, knowing is situated, specific to contexts and mediated by artefacts and distributed within the social environment, without any single individual holding all the relevant knowledge. Following this understanding, there is a commitment to participatory and action-oriented approaches. This follows the ascribed central role of the expertise of members of practice contexts, as well as the orientation to a region, which includes many actors, connected to each other in various ways. Via participatory approaches we aim to include these manifold perspectives and knowledges into the socio-technical interventions. This aim is embodied in the various activities and formats the Center offers (see section 4.2). All this, and this is crucial to note, requires a 'broad' stakeholder and context-sensitive view, to receive a more holistic picture of the problem situation in complex (socio-technical) systems (Ogonowski et al. (2013)). Such kind of studies are then to be understood as action research (Hayes (2011)), aiming to generate knowledge through intentional, directed intervention into practices. In such action research-oriented collaborations researchers learn from practice and vice versa (Baskerville (1999)). Even if such studies do not produce generalizable results (and do not intend to do so), they do result in an in-depth understanding of specific contextualized practices which provides a foundation for the design of IT artefacts (Rohde et al. (2017)).

Such a practice-oriented and participatory approach for HCI and CSCW research makes it a suitable framework for a technology transfer project as the one under investigation, where the primary goal of the project is not the creation of

scholarly knowledge but the creation of value for practice partners such as the regional rural SMEs in focus here. Again, from a practice-based perspective these formats and activities are located at the intersection between research and academia on the one hand and companies as well as other institutions on the other.

Activities of the Center

To realize these goals of regional knowledge co-creation based on a situated and practice-oriented approach the Center engages in a variety of activities for and with SMEs as well as other actors. Although the Center's program consists of a broad and diverse array of activities, they can broadly be summarized into the three categories of 1) Informational Events, 2) Workshops and 3) Implementation Projects. *Informational Events* are low-threshold activities, in which members of the center hold presentations about specific technologies or organise events where external experts speak, followed by an opportunity of attendees to ask questions. They can also take the form of a booth at an industry fair or the demonstration of a specific technology. They often are the first point of contact between the Center and SMEs. *Workshops* provide a more intense and hands-on introduction to specific topics. They are often combined into workshop series of 4-7 individual workshops, such as the digiXpert series.¹ Workshops usually center on real-world cases of the participant's companies. They focus on employees and the challenges they face in their work practices. *Implementation Projects* are cooperative design projects, in which a specific technology is prototypically deployed in a company context to create and opportunity to 'try it out'. These projects often build upon the challenges and ideas company employees brought to the workshop or voiced throughout. If no previous workshop experiences have been made to base design projects on, the specific topic to be addressed is decided jointly between company members and employees of the center, and subject to change after beginning the cooperation. Within small design projects these ideas are developed into socio-technical applications that can be applied and tested within the companies. Although these design projects usually do not take longer than a few months, they specifically follow a participatory and practice-based approach. This means that the experiences and perspectives of members of specific practice contexts take center stage, and any interventions start from ideas, challenges or wishes expressed by them. The primary goal of these design cases is not the development of large-scale systems to be applied within the company context, but to explore and illustrate IT-opportunities to support qualified work in a targeted manner and generate knowledge within the company to carry out their own digitization projects, in whatever form suits them.

¹ the name of the workshop series was changed for the purpose of this paper, to preserve anonymity as much as possible

Methodology - Doing Meta-Research

To describe our activities within the field, we conducted research on our own activities in the field (Dachtera et al. (2014)). As mentioned above, the Center of Competency under investigation began its work in late 2017. After an initial set-up phase, the activities described in the previous section began shortly after the inauguration of the project, within 2017. Overall, by 2021, four iterations of the digiXpert workshop series had been conducted, 28 design or implementation projects have taken place and members of the Center have organised or participated in 421 events. Since the beginning of the project, the researchers that are engaged in the project kept research diaries, taking short notes during specific events, steps and meetings which were later extended to detailed field notes (Argyris et al. (1985)). These include, for example, notes taken during the workshops of the digiXpert series, or transcripts of interviews with participants of the design projects as part of the collaborative investigation into specific practices in order to jointly plan design interventions. Secondly, qualitative open interviews were conducted with members of organizations that had participated in the activities of the Center, as well as members of the Center, as part of the internal evaluation of the Center's work. Such interviews are conducted periodically after the respective activities took place. Apart from identifying possible areas of improvement with regards to the program of the Center these interviews also served to gain a better, more detailed understanding of the measures of the Center, their perception by partners and participants, possible interdependencies and synergies. This evaluation is ongoing. At the current time it contains 21 interviews with a total length of 650 minutes. Nevertheless, it needs to be noted that data collection was somewhat opportunistic, as the funding organisation of the project explicitly rejects any engagement into 'research' by the employees of the center, and research activities therefore need to fit within the objectives of the center and its clients.

For the sake of this paper and the reflection of the Center's work that it entails, the data retrieved from the field notes as well as the interviews was combined. This data was then thematically analyzed (Braun and Clarke (2006)) in a collaborative manner. Authors compared and discussed their analysis and where necessary also sought clarification with members of participating organizations. Through this process, data was sorted into 'cases', with each case centering on one specific company, partnership or process and with a more or less central narrative.² Subsequently, a few cases were selected for this exploratory paper with the aim of providing preliminary insights into the work of the center, the regional focus and the associated hurdles. This process resulted in the narrative(s) presented in the section 'Case Studies'. To preserve anonymity, all names of actors, organizations, locations or events that could be potentially revealing have been altered. Names that are appear are pseudonyms

² Although it becomes clear throughout the following section that narratives are often connected and cases are not entirely distinct from each other, which we address in the discussion

Case Studies

In this section we describe four cases which illustrate the different forms the activities of the Center take and the different ways in which they manifest and intervene in the region. Case 1 describes an implementation project with a manufacturing company. Members of the Center worked with various employees of the company to develop a digital tool to support the work of the maintenance department. The resulting application later travelled to other departments and even to another company. Case 2 describes an implementation project in which members of the Center worked with an employee of a local company to develop an application to support quality management. Although never rolled out, for reasons we explain below, the co-design project served to create knowledge that enables the participants to make informed investment decisions about digitization efforts. Case 3 describes the evolving cooperation of the Center with an employee of the regional administration and her instrumental role in creating moments of knowledge co-creation. Case 4 describes how initial contact and exchange about digitization is established between various regional companies, and how these companies continue their exchange later without interference of the Center, co-constructing knowledge together.

Case 1: where an artefact is designed, put to work and then travels within the company

In the first case, a company worked with the Center of Competency on the co-design of an application to support the coordination of maintenance work within a specific department of the company. The system was applied in the company practice, appropriated, and even introduced by the company - without engagement of the employees of the Center - to other departments of the company. Later on, the system was also introduced to another company.

This co-design project was carried out with a medium-sized family-owned enterprise, typical for the companies in the region. Its business is the construction of appliances to equip factories, such as large cranes focusing on heavy duty lifting. This specific project addressed challenges regarding the organization of maintenance work within the company. The company has a small maintenance department, consisting of two employees, responsible for taking care of the machinery, tools, and facilities in a specific department of the company.

We began the project with a short inquiry into the maintenance work, via observations and conversations with maintainers and other staff over a few visits to the production site. These visits revealed that the department suffered from a lack of formal procedures to report maintenance issues and machine breakdowns, which created a variety of problems for carrying out the maintenance work. For example, issues would be reported to the maintainers unsystematically whenever staff spotted the maintenance workers in the construction hall. As maintainers do not always have documentation tools at hand, such sporadic reports were difficult

to keep track of. At other occasions little written notes or broken tools were simply left on their desk, without further information about the nature of the malfunction or the urgency of its repair. This made it difficult to keep track of issues and to prioritize the maintenance work appropriately, which affected both the maintainers work but crucially also production and resource planning. In the end the head of production often did not have a full overview of which machines were broken and to ultimately re-direct resources and orders to other machines and re-plan production accordingly. One maintainer explained this in the following:

“The broken tools are simply placed here on the table by the late shift just after they broke down. We find them in the morning. For example, we begin our work, and three small grinding tools are lying here on our desk, without any further information. [...] It could also be that someone tells you about a maintenance issue in passing and then you forget it. And when that happens perhaps twice, that we forget about an issue told to us in passing, at some point the topic gets taken to the foreman or superintendent and they then think that for six weeks the entire plant has not been run sensibly, as they were not aware about maintenance issues and breakdowns.”

We then went on to organize a series of design workshops at the company, involving employees of all hierarchy levels, including the maintainers and managers. The workshops served to discuss, further elaborate on and analyze what we had learned about the processes and the associated challenges so far and to begin to imagine applications to address them. We developed several concepts to illustrate how a digital artefact could support different practices of reporting and addressing maintenance issues. During the discussions that followed, the company staff decided on a mobile application that would support more structured means of reporting and prioritizing breakdowns and maintenance work, which was subsequently collaboratively designed and developed. Once a working prototype status had been reached, the application was introduced into the work context together, which required all employees to change their procedures regarding maintenance slightly - which had been the goal of the intervention. After this, we left the company, as the objective of the center to provide opportunities to test digital technologies in a prototypical manner, not to develop working solutions, had been fulfilled and the company also told us that we would not be needed anymore, after almost a year of working together.

A few weeks later we headed again to the factory to conduct a small evaluation of the application's use. The software was found to be still in use and supporting the cooperation of production and maintenance workers, to adapt new maintenance and reporting procedures. Instead of an illustrative prototype, the application had become an integral tool. We were furthermore surprised to learn that members of the company had decided by themselves to roll out the application in other production facilities. While this was exciting, it was also a bit unsettling, as we did not anticipate this and did nothing to support it. It happened outside of our control.

Lastly, the application also travelled to an entirely different company, whose members faced similar difficulties. Towards the end of the cooperation with the

first company a regional fair took place, focusing on digitization for SMEs. The Center was exhibiting their work and included a few technological demonstrators, such as the application to support maintenance work. At this fair, the CEO of another company approached the booth and became interested in this specific application. He was initially interested, but not convinced. During the following presentation a member of the company in which we first implemented the software joined the presenter and acted as a mediator and advocate, convincing the CEO of the second company by sharing his experience of the project carried out in his company. He mainly repeated the summary of the project given by member of the Center of Competency. His account focused on the work processes of the members of the maintenance department and how they changed due to the project, towards greater transparency and prioritization. He mentioned furthermore that both the process as well as the resulting application were received very well by the participating employees. Only then the CEO agreed to carry out a similar project in his company using a similar approach with regard to maintenance work. This design project is currently being carried out, focused on the adaptation of the system to the new environment and its appropriation.

Case 2: where an artefact is designed and without application enables informed future decisions for company members

In the second case we present, members of the Center collaborated with employees of a regional industrial company on the development of a tablet-run app to assist in quality management. In contrast to the company in the center of Case 1, the company of this case is a large company with about a thousand employees distributed across various production sites around the globe and the main site in the region in question. While the collaboration resulted in a fully working prototype of the app, the application was never rolled out. Towards the end of the development phase the company's IT department voiced strong security concerns and blocked its implementation. Nevertheless, to our surprise this was not perceived as a disappointment by our collaborators. In their view, the knowledge gained during the co-creation of the application was a sufficient benefit. It enabled them to understand their own needs and evaluate commercially available applications better, from which they finally picked a solution.

The project began through the involvement with the local office of the union of metal workers³. An employee of the company, who was responsible for quality management of specific products, learned about the program of the Center at a local union-organized event. In an interview conducted after the end of the project for evaluation purpose he told us:

"The metal union always organizes a 'market of possibilities', where they present what is new in the area of automation and digitization, but also new laws or legal decisions by judges. It's a kind of educational event that is very popular. And at the last edition I met the coordinator of the Center, and we talked about the

³ <https://www.igmetall.de/>

Center and all they can do – we talked for a very long time. I really liked what you presented."

After the event, the employee and the members of the Center stayed in contact via email, discussed challenges within the company and finally agreed to engage in a co-design project. The company employee and his supervisor, the head of the quality management department met with several members of the Center, to discuss which specific challenge to address and how they could be addressed. The company members were interested in reducing the amount of paper-based checklists in their assessment of the quality of specific large machine parts they manufactured. They had already very specific ideas what the tool they were setting out to design should entail: it should be a tablet-based app that would facilitate the completion of checklists to assess manufacturing quality and enable the inclusion of photos of the assessed part at specific items of the checklists. The inclusion of images was supposed to ease documentation of manufacturing faults.

The group of two company employees and three members of the Center now met regularly to work on the app. While the members of the Center took on the technical development, these meetings also included several visits to the shopfloor and discussions with other company employees, to gather further insights to inform the design. After several months a working prototype was ready, and the employee and his supervisor wanted to test it in practice with several further employees. To this end the company's IT department was contacted with a request for support, for example by supplying extra devices on which to run the application, who had so far been kept out of the project deliberately by the company members. The IT department however voiced concerns that the developed app violated some of their own security guidelines, such as running on the Android operating system instead of on Apple's iOS, which they exclusively used due to security issues. They blocked any further testing and development of the application. While this was of course an obvious source of possible disappointment, the supervisor did not think so - to our surprise: *"The idea was a good one, it gave us quite a lot of impetus - the work was in no way wasteful. It actually showed us quite clearly what we have and what we need. And we are now looking at what is available on the open market."* The project had seemingly enabled them to better understand their specific demands of a tool and now helped them make informed choices when evaluating commercially available applications.

Similarly, to the project of Case 1, the project of this second case was also not fully confined to the context in which it was developed and envisaged. During the development phase a member of the company mentioned the project in an article in the magazine published by the local chamber of commerce. He told us that he was subsequently called by a colleague from the sales department, who was interested in a similar check-list based tool but for a different use case, as they also complained about "too much paper". They met and discussed their possible options, but before this could develop any further the IT department put a halt to the entire project. After the end of the project, the application travelled further to an entirely different

context, the care sector, where it inspired further development. This is the subject of Case 3.

Case 3: where a partner initiates knowledge creation through her own networks, and an artefact travels from one sector to another

At the center of this case is the long-standing and evolving cooperation with an employee of the regional administration, which involves a variety of activities, from individual workshops and digiXpert workshop series to the co-creation of technological artefacts, which we will detail below. The application outlined in Case 2 also makes an appearance in this case.

A key partner of the Center is the regional development office of the county the Center is located in. Our key collaborator in that office, Ulrike, is both responsible for economic development as well as for the coordination of support for the health care sector in the region. The care sector is comprised of various organizations, including numerous smaller private care institutions which are SMEs. In this second capacity Ulrike organized an event on the "future of care" in the region, taking place in 2018, inviting various actors including care professionals from private and public institutions as well as technology providers and a member of our Center. Crucially, this event was initiated and organized by Ulrike, without involvement of the Center. She encouraged participation through her own network, inviting members of care companies as potentially interested participants via personal phone calls. Members of the Center then joined one of the workshops that were part of the event. While providing expertise on digitization in the workshop, their role was also to present the digiXpert workshop series, in order to generate interest for the series specifically within the care sector. Two employees of different care institutions signed up for the workshop, including Antje, a nurse at one of the participating care institutions, who was initially skeptical about the opportunities of digitization for her sector.

One goal of the workshop was to create an overview over available digital technologies that could support documentation needs within care practices. During the workshop it became clear that care professionals have a high load of documentation requirements, needing to, for example, regularly manually measure and document vital signs of patients several times a day. The member of the Center suggested that some of these tasks could be made easier by digital technologies and that the Center would be able to explore opportunities and maybe develop a prototype. This could build on technologies the Center had explored in industrial companies, such as an app to keep, manage and complete checklists, described in Case 2, which could be adapted for this context. Since this proposition was met with interest from several of the participants, one result of this workshop was a working group involving five regional care institutions and the Center, under the guidance of Ulrike and including Antje, the future ambassador.

The group began to meet regularly, albeit infrequently, coordinated by the Center member. In the first meeting the group was also joined by another member who had been a participant in one of the workshop series. In this meeting the

group decided that a smart watch might be a useful device to support the documentation tasks as it would eliminate the need to carry pen and paper around and digitize the data later. The smartwatch could include a checklist and enable fast and immediate recording of the data in digital form. To gather more information the group then organized visits to the respective care institutions. Here, it quickly became clear that a smart watch was not a feasible device to support documentation. Due to hygiene regulations, care practitioners in all institutions were not allowed to wear watches while working. A telephone would be more suitable device to run a checklist-like application, which was still considered the most appropriate way to address the documentation challenges. A first prototype was developed and discussed during another workshop, organized by Ulrike. During further visits to care facilities, however, the group discovered that many of the devices that were measuring patient-data and from which care practitioners copied the vital data manually, actually contained digital interfaces from which data could be taken automatically and sent somewhere else. This changed the task of the group considerably, as the goal was no longer the design of a checklist-app, but also the automatic incorporation of data from the devices that were able to send it. Ulrike then organized yet another workshop to narrow down the options and decide on one direction. During this final workshop the group sketched two further prototypes, in addition to the smartphone-based checklist-like app to facilitate digital documentation of patient data. These were 1) a hygiene-monitoring system that would access and gather data from the respective devices such as fridges, temperature- and humidity sensors in the care facilities to ensure proper hygiene or 2) the connection of a pulse-oximeter to continuously and automatically gather vital data of care patients. The covid-19 pandemic and the protective measures then forbid any further meetings, and the group has been somewhat dormant up to the writing of this account in mid 2021. Ulrike however carried the ideas and insights developed during this process also into another region, involving some of the group's members in the process, to spur similar developments elsewhere. She met at various points in time with leading members of regional chambers of commerce, unions as well as politicians from federal government to present and discuss these results and especially the benefits of the workshops series, for the care sector and beyond. Her goal was to inspire the development and uptake of such workshop series and participation of the care sector in other regions, even if the Center of Competency would not be involved, as she had witnessed the potentially beneficial role such workshops would pose for the sector. Similarly, Antje also began to present her experiences from the workshop series and the checklist app procedure at other opportunities, even in national events, but also regionally. Even though the process in the previous group lay dormant during the pandemic, Antje remained in contact with the Center, and initiated a cooperation with another care facility she was in contact with, who had previously not been involved in the group. In this new care facility, the previously sketched prototypes were put into action, and together several devices that measured environmental and health data, which care staff needed to record several

times a day, were connected to a locally hosted platform via their pre-existing interfaces. Staff then did not need to measure and collect this data manually, going around the facility, but the data was automatically sent to a dashboard which gives them access to an overview over all required data.

In the meantime, the cooperation with Ulrike was dormant for most of 2020, but began anew in early 2021, and is ongoing in various formats. As mentioned above, the core take-away of the case is not so much the design of technical artefacts and their effectiveness in practice, but the evolving cooperation with Ulrike and later on also with Antje. Both have been instrumental in facilitating the co-operation between the Center and regional companies and other actors in the care sector. They created connections through their own personal contacts, through which further projects with other actors evolved. Especially Ulrike stressed the importance of this cooperation: *"The two of us, Center and my office, playing together is crucial for the region. [...] It helps us and you and drives connections in the region. It has become a network and it is vital to carry it forward, as both sides can endlessly benefit from it."*

Case 4: where companies begin to create knowledge together

During one iteration of the digiXpert workshop series several participants expressed the wish to be in closer and more regular contact with other companies to exchange knowledge and learn together. It was important that companies should be similar in size and structure, for the experience to be more relatable and for needs, resources and constrictions to be more easily comparable, but should come from different sectors that do not pose any economic competition. This however was at first not easy to achieve. A participating employee from a rather large industrial company, in the team of production management, told us that he tried to engage with members of other companies he met at the workshop. *"A colleague from the workshop wanted to visit our company. And I said: yes, great! But I met him recently at an event and he told me that he just does not have time currently. They have similar processes [in their company]. We could both imagine to meet and exchange more often, but haven't managed due to a lack of time"*. Through the work of the Center members got to know several companies with interest in similar technologies, facing similar challenges within their production, but without direct competition between them, and organized regular meetings.

In these meetings various application scenarios were subsequently discussed, as well as problems that remained open, such as the digital administration of tools and the associated standards. After some time, these groups became self-organized by the participating company members. An employee of one of the participating companies shared with us: *"We are now coming together in a small group. There is [name of other company], who were also in exchange at that time and who are also active in the Centre [of Competency] and are always at the events, and the other company is directly in [name of the same city]. The advantage is that there is a shared interest in the set-up process of machines, and we are quite similar as*

companies. [...] And here we have found a regular exchange of ideas for us, we try to meet once or twice a year and then we meet also in smaller groups where we talk in a more technical manner." The participating companies then also realized benefits of the work groups beyond the initial shared interest or topic. They found that they are investigating similar topics and technologies and are exploring ways to cooperate and share knowledge and resources in these efforts: "We also want to try to bundle up a little, so that not everyone invests money in prototypes and everyone starts again from scratch to find out that we are trying to do this together."

Discussion and Lessons Learned

Previous studies from within the HCI community have investigated and reported on innovation ecosystems as context for HCI work, such as work by Avle, Lindtner or Freeman (see e.g. (Avle and Lindtner (2016); Avle et al. (2017); Lindtner et al. (2016, 2014); Freeman et al. (2019, 2018))). Such studies have largely been empirical in nature. They are crucial work for the HCI and CSCW community as they initiated the discourse and investigation of innovation ecosystems and regions as context for design research work. Our work intends to develop this discourse further by providing a report on an attempted intervention in such a regional innovation ecosystem. The cases described in the previous section exhibit a variety of ways in which the efforts of the Center of Competency initiate change in the region, affect SMEs and involve a variety of actors, including company employees and middle managers, university researchers, representatives of chambers of commerce and trade unions as well as members of public administration.

The experiences described in the cases above hold three preliminary lessons for CSCW research targeting regional innovation ecosystems. Firstly, the cases contain elements of knowledge co-creation through activities rooted in a practice-oriented approach, thereby hinting that such an approach holds benefits for what is usually referred to as 'transfer projects'. Secondly, it becomes clear that the effects of our interventions extend beyond the borders of collaborative research and design projects with individual companies, which are arguably a common form of practice-oriented CSCW projects. This thus presents somewhat of a methodological challenge for a practice-oriented approach, as participants and their relations extend significantly beyond the borders of what is usually understood as a specific practice context. The question emerges how CSCW research and design projects can take the complexity of contexts into account, and account for their own role in them. Thirdly, and following from the second lesson, a regional focus seems to require increased flexibility from university and research staff in what are considered appropriate activities for university members, in order to deal with the necessary relations and reputation, which can pose further significant challenges.

Below, we will outline and discuss each of these aspects individually and summarize their possible implications for HCI work with SMEs in regional

ecosystems. Lastly, we will outline what we perceive to be limitations of this approach.

Co-creating Knowledge

Although they entail different actors and activities, in all cases presented above collaborative knowledge creation emerges as a shared outcome. Companies and researchers learn the possibilities digital technologies hold for a specific company practice, members better able to formulate their own requirements, etc. This knowledge creation happens often through joint practical work, taking for example the form of short design projects, where a specific technological application is created or appropriated for a specific context, but also through collaborative activities during a workshop that address real-world challenges participating members of companies face.

It is thus crucial to notice that this does not represent 'knowledge transfer' even if that is part of the funder's objectives and brief, but knowledge co-creation. Knowledge is not transferred from the university or the Center of Competency to a company context, it is not taught or otherwise communicated through writing or video, but newly created for the challenges at hand through collaborative activities, starting with members of practice contexts such as specific production lines as for example in Case 1, instead of only management. These activities require that the different expertise(s) of the participating actors are brought together and combined in the practical collaboration. For example, while members of the Center might have expertise necessary to create digital technology, such as coding skills, graphic design etc., they do not know exactly how these skills might be useful in a specific context. Company members know this. Together an application or a prototype is developed. Co-creation requires both kinds of expertise (and more). This echoes an understanding of knowledge and learning formulated by Fischer et al. (Fischer et al. (2007)), as "mediated by artefacts, situated and distributed in a social environment" (p. 3), as we mentioned before, and at least tentatively signals that such an approach is able to build up knowledge within the companies, relevant to digitization. The approach of the Center also seems to present a divergence from how some of the companies in our cases so far have approached digitization, as for example shown in Case 2, and new methodological knowledge is co-created. Here, the IT department of the company was explicitly excluded by the participants from the company, to avoid their usual top-down approach and enable a different experience for the participating employees. In another sense, knowledge co-creation also takes place through the designed artefacts, that are appropriated beyond their initial context of application. For example, the activities described in Case 2 and 3, take an artefact designed for an industrial context to the care sector, where it finds another use. It does however not simply travel to be applied, but it is collaboratively and actively altered, appropriated, via the working group that emerged during the initial workshop, to make sense and be of use in the intended new context of the care sector. Similarly, knowledge travels within the company in

Case 1, where the artefact is appropriated into another production line of the same company, or in Case 2, where it at least inspires a colleague from a separate department to engage in further digitization projects. Especially Case 1 therefore illustrates that a practice-oriented and participatory approach is suitable to knowledge co-creation, as the co-workers at the production site were able to transfer and adapt the application to another production line without involvement of the researchers - something which was not anticipated or explicitly planned for.

However, that the effects are not solely bound to a single location, but somehow 'travel' through the region in different ways is both beneficial as well as problematic for our chosen approach. We will reflect on the travelling nature of effects and artefacts in the next section.

Effects travel through the ecosystem

The cases above also illustrate, that the consequences or effects of such work are not bound or limited to a specific place or a specific group of actors. The Center of Competency acts not only within a specific department or with a defined group of people, but targets the entire region, comprised of various cities and villages and various organizations. The region is then rather a complex network of actors, and actions within such a complex network have unforeseen consequences and effects "travel". Several cases illustrate this phenomenon of travelling effects. In Case 4, for example, the members of the Center worked with a specific, even rather small group of people in a few companies as participants in the digiXpert workshops. Yet after initiating exchange between these specific members of companies, the constellation of companies involved in the co-creation process changed, without the Center's interference, self-directed by the needs and interests of the company employees, and the circle of involved companies even grew. Similarly, in Case 1 the maintenance system is applied in other parts of the company, outside of the sphere of the Center, taking a more direct way than the newspaper. In Case 2 and 3, the effects travel between sectors, from the manufacturing of heavy machinery to the care sector. In Case 3 co-creation is initiated and continued by our collaborator from public administration with other stakeholders, building on and engaging their own personal and professional relationships, outside the realm of the Center of Competency, and later the Center of Competency is involved again.

These forms of travelling of artefacts and their effects can be understood as form of spillover. Spillover is a known phenomenon in other disciplines, such as economics and management studies (see e.g. (Mascarenhas et al. (2018); Scarrà and Piccaluga (2020); Alcácer and Chung (2007); Audretsch and Feldman (2004)) yet relatively unexplored in CSCW and wider HCI. Spillover also seems to create some problems for a practice-oriented form of research and design, as we will discuss. Even though the work of the Center strongly builds on the practice paradigm outlined above, the ways in which this work of the Center of Competency takes hold in the region expands the usual frame of research and design projects within this practice paradigm. While this is not necessarily a

diversion from the epistemological orientation towards practices, it presents a methodological challenge of how to address various, connected practices or practice contexts. The activities and effects transcend the localized nature of situated practices and become distributed across the network of actors of the region instead. Design and knowledge co-creation might initiate in the Center's activities, but continues in locations and through relations of which the members of the Center are not only not a part, but of which existence they might not even be aware. This makes it difficult to remain accountable for the effects of one's activities. Lucy Suchman has reflected on this phenomenon in her article "located accountabilities" (Suchman (1993, 2002)). In this work she shows how technology design, as the materialization of knowledge, is not located within the distinct group of designers, but distributed across various actors, where 'use' is an active, creative act of appropriating technologies, giving them meaning and making them fit. Taking this distributed nature serious, means accepting "the limited power of any actors or artifacts to control technology production/use;" (Suchman (2002)). The challenge thus becomes how to engineer for this kind of spillover from the initial site and context where a design project was carried out, facilitating the travel across various contexts, while also acknowledging that one has limited control or even access to the relationships that make spillover possible. A possible opening for such engineering for spillover could lie in the routinised nature of practices, and that such routines can be similar from one context to another. Identifying similar routinised practices in different context allows thus to apply similar technologies. At the same time, this only partially explains the travels beyond single contexts we describe here, as such travelling can be witnessed when similarities between practices seem absent. Further research is required that pays attention to the structures and practices involved in spillover in order to formulate approaches that might facilitate it.

A further challenge thus arises, that, in order to engineer for spillover, we need to access or build the relationships through which it takes place. In our experience, this requires activities that are often significantly different from what is usually considered research or design.

Necessary Work Beyond Research and Design

Apart from the activities outlined above in the section titled 'Case Studies', which constitute perhaps the core of the Center's work, members are also engaged in constant activities that constitute the building and maintenance of networks, relationships and reputation. Activities are carried out constantly to keep connections with partners alive and build new ones. Such activities include (of course) countless meetings, email exchanges and telephone calls to actively plan projects or simply engage in relationship building for future joint activities (see experiences from other scholars, such as (Ahmadi et al. (2020); Meurer et al. (2018)). Some of these activities are portrayed in what is above called "Informational Events", which can include presentations at events such as the

market of ideas mentioned in Case 2, the fair in Case 1, where the members of the two companies met as well as visits to companies and other organizations. Not all activities, however, are included in such work packages, and some activities thereby take place outside the project, at least in a formal sense. Through such activities however, relations are maintained and created which later become the foundation for the co-construction of knowledge, or at least the starting point for co-construction processes, and for which one can account as a HCI or CSCW designer-researcher. The point is, however, that these activities are not what is typically considered research or design. They nevertheless take up a central role in such region-oriented projects, more, we would argue, than in projects focused on a specific and more confined or bounded practice context, where fewer relations are involved. For a university-based project this constitutes some challenges, as it is not a trivial undertaking to translate such activities into what is usually thought of as research outputs, including the preparation of research publications or other academic activities such as teaching of students and other academic qualifications such as dissertations. Others, such as Meurer et al. (2018) or Krüger et al. (2021) have also pointed out the need for intensive networking activities as the foundation for participatory research activities, and especially to make such efforts last. This is reminiscent of work by Blomberg and Karasti (2013), who have highlighted the fact that 'field sites' are always designed. While we do not mean to highlight here the designed and thereby artificial nature of any field site, which is certainly also the true for our work, constructing a field site - for research and for design - takes work. In the case of the Center of Competency, this is not just a single field site, but multiple sites, across companies and departments. This is especially urgent, we believe, as the regional focus requires a constant creation and maintenance of relationships and reputation, which do not look anything like 'co-design', but create the foundation for future design or research activities.

Conclusion

To conclude, these three lessons begin to paint a picture of what we might call 'engineering for spillover'. Practice-oriented, collaborative research and design activities seem to be suitable to facilitate the co-creation of knowledge that enables SMEs to engage in digitization projects themselves. These effects of these activities and the artefacts that result from them are not always bound to specific contexts and locations but spill over into other departments, companies, sectors. This happens through connections, networks and relations, which one might or not might not be part of. Ultimately, addressing a region through spillover means accepting Suchman's lesson on the located accountabilities of technology designers that we have limited power to control technology use (Suchman (2002)). Nevertheless, it seems as if engineering for spillover is a possible way to support the wider effects of localized co-creation activities. This, we propose, includes a number of network and reputation-building activities which do not always look like research or design and are possibly quite mundane, but are nevertheless the

foundation for co-creation activities and their regional spillover effects. More research is however required to better understand the processes of spillover effects, the actors involved and their potential practices, in order to better engineer for spillover.

Acknowledgments

We would like to thank all our partners for their fruitful collaboration over the years. Thanks also to Peter Tolmie and Dave Randall for their invaluable feedback and criticism. The same gratefulness is extended to the reviewers and their comments and recommendations. This work is partially funded by the Federal Ministry of Education and Research (BMBF), grant number 16SV8788.

References

- Ackerman, M. S. (2000): 'The Intellectual Challenge of CSCW: The Gap Between Social Requirements and Technical Feasibility'. *Human-Computer Interaction*, vol. 15, no. 2-3, pp. 179–203. Publisher: Taylor & Francis _eprint: https://doi.org/10.1207/S15327051HCI1523_5.
- Adams, S. B. (2005): 'Stanford and Silicon Valley: Lessons on Becoming a High-Tech Region'. *California Management Review*, vol. 48, no. 1, pp. 29–51.
- Ahmadi, M., R. Eilert, A. Weibert, V. Wulf, and N. Marsden (2020): 'Feminist Living Labs as Research Infrastructures for HCI: The Case of a Video Game Company'. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 1–15, Association for Computing Machinery.
- Alcácer, J. and W. Chung (2007): 'Location Strategies and Knowledge Spillovers'. *Management Science*, vol. 53, no. 5, pp. 760–776.
- Argyris, C., R. Putman, R. D. Putnam, D. M. Smith, and others (1985): *Action science*, Vol. 13. Jossey-bass.
- Asheim, B. T. (1998): 'Territoriality and economics: on the substantial contribution of economic geography'. *Svensk Geografisk Årsbok*, vol. 74, pp. 98–109.
- Asheim, B. T. and L. Coenen (2005): 'Knowledge bases and regional innovation systems: Comparing Nordic clusters'. *Research Policy*, vol. 34, no. 8, pp. 1173–1190.
- Audretsch, D. B. and M. P. Feldman (2004): 'Chapter 61 - Knowledge Spillovers and the Geography of Innovation'. In: J. V. Henderson and J.-F. Thisse (eds.): *Handbook of Regional and Urban Economics*, Vol. 4 of *Cities and Geography*. Elsevier, pp. 2713–2739.
- Avle, S. and S. Lindtner (2016): 'Design(ing) 'Here' and 'There': Tech Entrepreneurs, Global Markets, and Reflexivity in Design Processes'. pp. 2233–2245, ACM Press.
- Avle, S., S. Lindtner, and K. Williams (2017): 'How Methods Make Designers'. pp. 472–483, ACM Press.
- Baskerville, R. L. (1999): 'Investigating Information Systems with Action Research'. *Communications of the Association for Information Systems*, vol. 2.

- Bergenholtz, C. and C. Waldstrøm (2011): 'Inter-Organizational Network Studies—A Literature Review'. *Industry and Innovation*, vol. 18, no. 6, pp. 539–562. Publisher: Routledge _eprint: <https://doi.org/10.1080/13662716.2011.591966>.
- Bjørn, P., L. Ciolfi, M. Ackerman, G. Fitzpatrick, and V. Wulf (2016): 'Practice-based CSCW Research: ECSCW bridging across the Atlantic'. In: *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion*. New York, NY, USA, pp. 210–220, Association for Computing Machinery.
- Blomberg, J. and H. Karasti (2013): 'Reflections on 25 years of ethnography in CSCW'. *Computer supported cooperative work (CSCW)*, vol. 22, no. 4, pp. 373–423.
- Braun, V. and V. Clarke (2006): 'Using thematic analysis in psychology'. *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101.
- Bundesministerium und für Ernährung und Landwirtschaft (BMEL) (2015): 'Ländliche Regionen verstehen · Fakten und Hintergründe zum Leben und Arbeiten in ländlichen Regionen', pp. 40.
- Chesbrough, H., W. Vanhaverbeke, and J. West (2006): *Open Innovation: Researching a New Paradigm*. OUP Oxford.
- Csikszentmihalyi, C., J. Mukundane, G. F. Rodrigues, D. Mwesigwa, and M. Kasprzak (2018): 'The Space of Possibilities: Political Economies of Technology Innovation in Sub-Saharan Africa'. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. p. 306, ACM.
- Dachter, J., D. Randall, and V. Wulf (2014): 'Research on research: design research at the margins: academia, industry and end-users'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 713–722, Association for Computing Machinery.
- de Castro Leal, D., M. Krüger, V. T. E. Teles, C. A. T. E. Teles, D. M. Cardoso, D. Randall, and V. Wulf (2021): 'Digital Technology at the Edge of Capitalism: Experiences from the Brazilian Amazon Rainforest'. *ACM Transactions on Computer-Human Interaction*, vol. 28, no. 3, pp. 18:1–18:39.
- Fischer, G. (2000): 'Symmetry of ignorance, social creativity, and meta-design'. *Knowledge-Based Systems*, vol. 13, no. 7, pp. 527–537.
- Fischer, G., M. Rohde, and V. Wulf (2007): 'Community-based learning: The core competency of residential, research-based universities'. *International Journal of Computer-Supported Collaborative Learning*, vol. 2, no. 1, pp. 9–40.
- Fowles, R. A. (2000): 'Symmetry in Design Participation in the Built Environment: Experiences and Insights from Education and Practice'. In: S. A. R. Scrivener, L. J. Ball, and A. Woodcock (eds.): *Collaborative Design*. London, pp. 59–70, Springer.
- Freeman, G., J. Bardzell, and S. Bardzell (2019): 'Open Source, Open Vision: The MakerPro Network and the Broadening of Participation in Setting Taiwan's IT Vision Agenda'. *Human-Computer Interaction*, vol. 34, no. 5-6, pp. 506–540. Publisher: Taylor & Francis _eprint: <https://doi.org/10.1080/07370024.2018.1555043>.
- Freeman, G., S. Bardzell, and J. Bardzell (2018): 'Bottom-Up Imaginaries: The Cultural-Technical Practice of Inventing Regional Advantage through IT R&D'. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. Montreal QC, Canada, pp. 1–11, ACM Press.

- Følstad, A., P. B. Brandtzæg, J. Gulliksen, M. Börjeson, and P. Näkki (2009): 'Towards a Manifesto for Living Lab Co-creation'. In: T. Gross, J. Gulliksen, P. Kotzé, L. Oestreicher, P. Palanque, R. O. Prates, and M. Winckler (eds.): *Human-Computer Interaction – INTERACT 2009*. Berlin, Heidelberg, pp. 979–980, Springer.
- Grant, R. M. and C. Baden-Fuller (2004): 'A Knowledge Accessing Theory of Strategic Alliances'. *Journal of Management Studies*, vol. 41, no. 1, pp. 61–84.
- Gulati, R. (1998): 'Alliances and networks'. *Strategic Management Journal*, vol. 19, no. 4, pp. 293–317.
- Hagedoorn, J. (1993): 'Understanding the rationale of strategic technology partnering: Interorganizational modes of cooperation and sectoral differences'. *Strategic Management Journal*, vol. 14, no. 5, pp. 371–385.
- Haipeter, T. (2012): 'Sozialpartnerschaft in und nach der Krise: Entwicklungen und Perspektiven'. *Industrielle Beziehungen. Zeitschrift für Arbeit, Organisation und Management*, no. 4, pp. 387–411.
- Hamel, G. (2011): 'Competition for Competence and Inter-Partner Learning within International Strategic Alliances'. In: *Organizational Collaboration*. Routledge.
- Hardy, J., D. Dailey, S. Wyche, and N. M. Su (2018): 'Rural Computing: Beyond Access and Infrastructure'. In: *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing*. New York, NY, USA, pp. 463–470, Association for Computing Machinery.
- Hardy, J., C. Phelan, M. Vigil-Hayes, N. M. Su, S. Wyche, and P. Sengers (2019a): 'Designing from the rural'. *Interactions*, vol. 26, no. 4, pp. 37–41.
- Hardy, J., S. Wyche, and T. Veinot (2019b): 'Rural HCI Research: Definitions, Distinctions, Methods, and Opportunities'. *Proceedings of the ACM on Human-Computer Interaction*, vol. 3, no. CSCW, pp. 1–33.
- Hayes, G. R. (2011): 'The relationship of action research to human-computer interaction'. *ACM Transactions on Computer-Human Interaction*, vol. 18, no. 3.
- Irani, L. (2019): *Chasing innovation: Making entrepreneurial citizens in modern India*, Princeton studies in culture and technology. Princeton, New Jersey ; Oxford, United Kingdom: Princeton University Press.
- Irani, L., J. Vertesi, P. Dourish, K. Philip, and R. E. Grinter (2010): 'Postcolonial computing: A lens on design and development'. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. pp. 1311–1320, ACM.
- Krüger, M., A. Weibert, D. d. C. Leal, D. Randall, and V. Wulf (2021): 'It Takes More Than One Hand to Clap: On the Role of Care in Maintaining Design Results'. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 1–14, Association for Computing Machinery.
- Kudic, M., M. Müller, T. Buchmann, A. Pyka, and J. Günther (2021): 'Network dynamics, economic transition, and policy design—an introduction'. *Review of Evolutionary Political Economy*, vol. 2, no. 1, pp. 1–8.
- Kuutti, K. and L. J. Bannon (2014): 'The turn to practice in HCI: towards a research agenda'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Toronto Ontario Canada, pp. 3543–3552, ACM.

- Lanamäki, A. and K. Väyrynen (2016): 'Six Issues in Which IS and CSCW Research Communities Differ'. In: A. De Angeli, L. Bannon, P. Marti, and S. Bordin (eds.): *COOP 2016: Proceedings of the 12th International Conference on the Design of Cooperative Systems, 23-27 May 2016, Trento, Italy*. Cham, pp. 3–19, Springer International Publishing.
- Lave, J. and E. Wenger (1991): *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.
- Lindtner, S., S. Bardzell, and J. Bardzell (2016): 'Reconstituting the Utopian Vision of Making: HCI After Technosolutionism'. pp. 1390–1402, ACM Press.
- Lindtner, S., G. D. Hertz, and P. Dourish (2014): 'Emerging sites of HCI innovation: hackerspaces, hardware startups & incubators'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 439–448, Association for Computing Machinery.
- Ludwig, T., C. Kotthaus, M. Stein, H. Durt, C. Kurz, J. Wenz, T. Doublet, M. Becker, V. Pipek, and V. Wulf (2016): 'Arbeiten im Mittelstand 4.0 – KMU im Spannungsfeld des digitalen Wandels'. *HMD Praxis der Wirtschaftsinformatik*, vol. 53, no. 1, pp. 71–86.
- Makoto Su, N., J. Hardy, M. Vigil-Hayes, T. Veinot, and R. Comber (2021): 'Introduction: Performing Rurality with Computing'. *ACM Transactions on Computer-Human Interaction*, vol. 28, no. 3, pp. 16e:1–16e:13.
- Mascarenhas, C., J. J. Ferreira, and C. Marques (2018): 'University–industry cooperation: A systematic literature review and research agenda'. *Science and Public Policy*, vol. 45, no. 5, pp. 708–718.
- Meurer, J., C. Müller, C. Simone, I. Wagner, and V. Wulf (2018): 'Designing for Sustainability: Key Issues of ICT Projects for Ageing at Home'. *Computer Supported Cooperative Work (CSCW)*, vol. 27, no. 3, pp. 495–537.
- Mowery, D. C., J. E. Oxley, and B. S. Silverman (1996): 'Strategic alliances and interfirm knowledge transfer'. *Strategic Management Journal*, vol. 17, no. S2, pp. 77–91.
- Nowotny, H., P. B. Scott, and M. T. Gibbons (2013): *Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty*. John Wiley & Sons.
- Ogonowski, C., B. Ley, J. Hess, L. Wan, and V. Wulf (2013): 'Designing for the living room: long-term user involvement in a living lab'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 1539–1548, Association for Computing Machinery.
- Ozman, M. (2009): 'Inter-firm networks and innovation: a survey of literature'. *Economics of Innovation and New Technology*, vol. 18, no. 1, pp. 39–67. Publisher: Routledge _eprint: <https://doi.org/10.1080/10438590701660095>.
- Pipek, V. and V. Wulf (2009): 'Infrastructuring: Toward an integrated perspective on the design and use of information technology'. *Journal of the Association for Information Systems*, vol. 10, no. 5, pp. 1.
- Pyka, A., M. Kudic, and M. Müller (2019): 'Entry dynamics of entrepreneurial firms, knowledge diffusion efficiency and regional innovation processes'. *Regional Studies*, vol. 53, no. 9, pp. 1321–1332.

- Randall, D., J. Dachtera, T. Dyrks, B. Nett, V. Pipek, L. Ramirez, G. Stevens, I. Wagner, and V. Wulf (2018): 'Research into design-research practices: Supporting—an agenda towards self-reflectivity and transferability'. In: V. Wulf, V. Pipek, D. Randall, M. Rohde, K. Schmidt, and G. Stevens (eds.): *Socio-informatics*. New York, NY, USA: Oxford University Press.
- Rittel, H. W. J. (1984): 'Second-generation Design Methods'. In: N. Cross (ed.): *Developments in design methodology*. Chichester ; New York: Wiley.
- Rohde, M., P. Brödner, G. Stevens, M. Betz, and V. Wulf (2017): 'Grounded Design – a praxeological IS research perspective'. *Journal of Information Technology*, vol. 32, no. 2, pp. 163–179.
- Saxenian, A. (1994): *Regional advantage: culture and competition in Silicon Valley and Route 128*. Cambridge, Mass: Harvard University Press.
- Scarrà, D. and A. Piccaluga (2020): 'The impact of technology transfer and knowledge spillover from Big Science: a literature review'. *Technovation*, pp. 102165.
- Schmidt, K. and L. Bannon (2013): 'Constructing CSCW: The First Quarter Century'. *Computer Supported Cooperative Work (CSCW)*, vol. 22, no. 4, pp. 345–372.
- Stein, M. (2017): 'From directions to actions—IT support for individual mobility in everyday activities'.
- Stevens, G., V. Pipek, and V. Wulf (2009): 'Appropriation Infrastructure: Supporting the Design of Usages'. In: V. Pipek, M. B. Rosson, B. de Ruyter, and V. Wulf (eds.): *End-User Development*. Berlin, Heidelberg, pp. 50–69, Springer.
- Stevens, G., V. Pipek, and V. Wulf (2010): 'Appropriation Infrastructure: Mediating Appropriation and Production Work'. *Journal of Organizational and End User Computing (JOEUC)*, vol. 22, no. 2, pp. 58–81. ISBN: 9782010040108 Publisher: IGI Global.
- Stevens, G., M. Rohde, M. Korn, and V. Wulf, 'Grounded Design - A Research Paradigm in Practice-Based Computing'. In: *Socio-informatics*. New York, NY: Oxford University Press.
- Strüker, J., M. Weibelzahl, M.-F. Körner, A. Kießling, A. Franke-Sluijk, and M. Hermann (2021): 'Decarbonisation through digitalisation : Proposals for Transforming the Energy Sector', pp. 19 pages. Artwork Size: 19 pages Medium: application/pdf Publisher: University of Bayreuth.
- Suchman, L. (1993): 'Working relations of technology production and use'. *Computer Supported Cooperative Work*, vol. 2, no. 1-2, pp. 21–39.
- Suchman, L. (2002): 'Located accountabilities in technology production'. vol. 14, pp. 16.
- Swiaczny, F. (2015): 'Auswirkungen des demographischen Wandels auf die regionale Bevölkerungsdynamik in Deutschland'. *Raumforschung und Raumordnung | Spatial Research and Planning*, vol. 73, no. 6, pp. 407–421.
- Tajnai, C. (1996): 'From the Valley of Heart's Delight to the Silicon Valley: A Study of Stanford University's Role in the Transformation', pp. 35.
- Vogelgesang, W., J. Kopp, R. Jacob, and A. Hahn (2018): *Stadt – Land – Fluss*. Wiesbaden: Springer Fachmedien Wiesbaden.

- Wulf, V., C. Müller, V. Pipek, D. Randall, M. Rohde, and G. Stevens (2015): 'Practice-Based Computing: Empirically Grounded Conceptualizations Derived from Design Case Studies'. In: V. Wulf, K. Schmidt, and D. Randall (eds.): *Designing Socially Embedded Technologies in the Real-World*. London: Springer London, pp. 111–150.
- Wulf, V., V. Pipek, D. Randall, M. Rohde, K. Schmidt, and G. Stevens (eds.) (2018): *Socio-informatics*. New York, NY: Oxford University Press.
- Wulf, V., M. Rohde, V. Pipek, and G. Stevens (2011): 'Engaging with practices: design case studies as a research framework in CSCW'. In: *Proceedings of the ACM 2011 conference on Computer supported cooperative work*. New York, NY, USA, pp. 505–512, Association for Computing Machinery.
- Ziman, J. (2000): 'Postacademic Science : Constructing Knowledge with Networks and Norms : Royal Society Medawar Lecture, 29 June 1995'. , vol. 9, pp. 93–113.

Nora O. Ringdal and Babak A. Farshchian (2022): The Role of Boundary Objects in Platformization Practices: A Case Study of Software Testing. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_ep07

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_ep07

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

The Role of Boundary Objects in Platformization Practices: A Case Study of Software Testing

Nora O. Ringdal and Babak A. Farshchian

Norwegian University of Science and Technology (NTNU), Trondheim, Norway

nora.o.ringdal@ntnu.no; babak.farshchian@ntnu.no

Abstract. While digital platforms are frequently investigated at technical, societal, and organizational levels, there are relatively few empirical studies of the collaborative practices that are involved when platforms are introduced into organizations. In this paper, we investigate such practices in the context of a large-scale platformization project within healthcare. Because off-the-shelf platforms already possess a stable core, platformization processes often focus on downstream system development activities such as configuration and testing. Our case study is about one such downstream activity, i.e., that of software testing. We frame software testing as a sociotechnical process involving tacit knowledge from a variety of user groups. We use the theoretical framework of boundary objects to demonstrate how test artifacts – mainly the test versions of the product – function as boundary objects, used to transfer knowledge among platform developers, those who configure the platform, and future users of the platform. Our findings show when and how boundary objects function or break down. We discuss the theoretical and practical implications of our findings both with respect to the boundary objects themselves, and the practices that surround boundary objects to support their collaborative properties.

Introduction

I feel like if I had something to say, I should have said it earlier

The above sentence is by a domain expert we interviewed during our study of a platformization process, i.e., the process of introducing a digital platform in her workplace. *Digital platforms* (Tiwana, 2014) are not only transforming our societies within transportation, tourism, education, finance, healthcare, etc. These platforms are also fundamentally altering our everyday work practices by, e.g., redefining the workplace itself (Ajzen and Taskin, 2021), outsourcing human resource-related decisions to algorithms (Jarrahi et al., 2021), and even redefining our organizational models (Faraj and Pachidi, 2021). With such a broad area of impact, it is difficult to define what a platform is. From a global perspective, platforms overlap with digital infrastructures such as the Internet (Plantin et al., 2018). From an organizational perspective, Gawer (2014) differentiates between *industry* platforms (such as the global platforms of Google and Facebook), *supply chain* platforms (with contractual relations among supply chain members), and *internal* platforms (developed within an organization for internal use). A common denominator for all platform definitions is that a platform consists of a *core* and its *periphery*, which creates a generative power that allows platforms to gradually adapt to different settings (Rodon Modol and Eaton, 2021; Tiwana, 2014). In this paper, we want to focus on a supply chain platform – in the form of an off-the-shelf product – and how it is being implemented in and reshaping an organization, here, a healthcare organization.

This process of *platformization* – which was the setting for the opening quotation from our domain expert in the healthcare organization – is multi-faceted. It is, for instance, studied in depth from a societal perspective as a process of “platform imperialism” driven mainly by American platform companies and their impact on various “spheres of life” such as social and cultural (see e.g., Poell et al., 2019). It is also studied from a technological perspective by, e.g., Bygstad and Hanseth (2019), defined as the process of moving from technological silos to platform-oriented infrastructures that connect the silos and facilitate data sharing. In this paper, we look at platformization *practices*, in particular when supply chain platforms are implemented in organizations. We define platformization as “the process of organizational, social, financial and technological transformation that an organization often must go through to effectively utilize a platform model” (Farshchian et al., 2021). Platformization is, therefore, a process that requires stakeholders to cooperate across professional, social, and/or cultural boundaries. Our case study represents snapshots from such a process, told from the perspectives of those users who are involved in this transforming process.

We have chosen boundary objects (Star and Griesemer, 1989) as our theoretical framework. Boundary object (BO) theory has been used in earlier research to describe system development activities (Doolin and McLeod, 2012) and platformization (Islind et al., 2019), and our study adds to this growing body of literature. As we will demonstrate, BO is a useful concept to study some of the

challenges in platformization processes where there is a strong and established platform core acting as both a liberating and constraining force. This is true for our case, where the platform core is an established product purchased from a world-leading vendor. Framing this core as a BO contrasts the view that BOs should be malleable and negotiable in order to be useful (see e.g., Lee, 2007). We show that boundary objects – particularly the product and its graphical user interface – play a central role as educational objects (Akkerman and Bakker, 2011) and as “wake-up calls” for users to see what is coming. The platform core is therefore useful as a “non-negotiable” BO not only for knowledge transfer purposes but also for motivating users (an effect also observed by Ellingsen and Hertzum, 2019). Our findings also complement the views taken by Isind et al. (2019) and partly by Doolin and McLeod (2012), both studying the design of new systems. Our case – being about an existing system designed by others a long time ago – demonstrates the challenges posed by a non-negotiable or “closed” (Isind et al., 2019) BO during downstream design activities.

Our study also demonstrates the strengths and weaknesses of the BO concept in large-scale system development projects consisting of several organizations and hundreds of participants, where formal and top-down project management regimes can create new settings for a BO. This setting is similar to earlier studies of enterprise resource planning (ERP) software (e.g., SAP, see Schreieck et al., 2021) or global packaged software in general (Pollock et al., 2007). Our research builds on this body of literature and focuses on the practices of user participants during the downstream design activity of testing. There are few studies of testing practices in such settings. We demonstrate how testing becomes a central activity in platformization because – as formulated by our domain expert – many design decisions are already taken a long time ago to build the core of the platform (see also Williams and Pollock, 2012). This means that downstream activities such as testing and training play a potentially important role in platformization because they become the only avenue for users to participate in design. Our findings show that boundary objects – mainly in the form of the software product itself but also test plans, project plans, presentations, etc. – take center stage in testing. Our findings also demonstrate the challenge of cooperating with an international platform provider who becomes a permanent service provider and a powerful part of the ecosystem.

In this paper, we pose the question: What happens in downstream design activities in platformization processes? Our empirical data are from a large ongoing platform implementation project in Central Norway. The project has a complex ecosystem consisting of vendors, customers, and intermediaries, in this way implementing what Gawer (2014) calls a supply chain platform. The findings presented in this paper are preliminary and will be refined and updated as we collaborate with the case organizations and collect and analyze more data.

In the rest of this paper, we first discuss some relevant literature in the areas of platformization, boundary objects, and testing. We then provide a short description

of our case and the case study method we are using. We will then describe some of our findings and discuss their implications for theory and practice.

Background

Platformization

A digital platform (also called a platform in this paper) is defined as a “software-based product or service that serves as a foundation on which outside parties can build complementary products or services” (Tiwana, 2014, p. 5). The “foundation,” commonly called the core of the platform, is shared across all the users of the platform and is often difficult to change, i.e., it is “entrenched” (Rodon Modol and Eaton, 2021). The complementary products or services – i.e., the “generative” parts of the platform ecosystem (Thomas and Tee, 2021) – are usually referred to as periphery and are connected to the core using boundary resources (Ghazawneh and Henfridsson, 2013) such as application programming interfaces (APIs). As an example from our case study, the healthcare platform in discussion has a core consisting of an electronic health record (EHR) and a basic set of applications and APIs. This core is then configured and further developed with functionality – i.e., the periphery – needed to support the practices in each specific healthcare organization. Although platformization is a broad concept and contains transformations at different levels and scopes (Farshchian et al., 2021), configuring and augmenting the functionality of the technical platform is a central part of this process and is the focus of our study here. Configuring the platform to fit into the organization can be a challenge for several reasons. First, the organization will need to change to fit to the core functionality of the platform software (Strong and Volkoff, 2010). Second, the organization needs to take advantage of the generativity of the platform’s periphery to create innovations in its own context. This is a challenge because “While configuration makes some functionalities easy to set up, it also restricts the space of possible functionalities to those envisioned by the designers of the configuration facility” (Ellingsen and Hertzum, 2019, p. 2). Moreover, platformization is not a linear process with design following use. Due to their broad areas of impact, using platforms leads to organizational change that, in turn, requires reconfiguration of the technological base (Leonardi, 2009).

According to Pollock et al. (2007, p. 1) generic systems such as platforms are “brought into being through an intricately managed process, involving the broader extension of a particularised software application and, at the same time, the management of the user community attached to that solution.” The majority of existing studies have investigated the process of platformization on technical, societal, and organizational levels. Fewer studies have explored the collaborative practices that are involved when platforms are introduced into organizations (see e.g., Ellingsen and Hertzum, 2019; Farshchian and Thomassen, 2019; Islind et al.,

2019). Our study builds on these and extends them by looking at downstream system development activities of testing

Software testing as a sociotechnical process

According to Myers et al. (2012, p. 2), software testing is a “process, or a series of processes, designed to make sure computer code does what it was designed to do and, conversely, that it does not do anything unintended.” According to some studies, testing can take up to fifty percent of the total development costs (Bertolino, 2007). Numerous classifications of software testing exist. The widely used textbook by Sommerville (2016) defines three major types of testing. *Development testing* is testing done by a system’s developers during its initial development to discover bugs and defects. *Release testing*, often performed by a separate group than developers, is about making sure that the system as a whole meets the requirements of the system stakeholders. *User testing*, including acceptance testing, is done by end-users or potential end-users in their own environment to document whether the system is useful and usable, i.e., whether its features are understood and accessible by its end-users and whether it supports end-users to do their job. Recent developments in software engineering, such as agile and continuous development, have blurred the distinctions between these types of tests. This is in particular true for release and user testing, which often happen in parallel.

While the vast majority of software testing research is concerned with the vision of fully automated testing (see e.g., Bertolino, 2007; Orso and Rothmel, 2014), software testing is also regarded as a fuzzy concept, dependent on the sociotechnical surroundings of a software project (Rooksby et al., 2009). This fuzziness seems to increase when testing modern types of software - e.g., platforms and mobile systems - characterized by high levels of heterogeneity and configurability Sommerville (2008). As noted by Orso and Rothmel (2014, p. 125): “testing of these systems is often performed in ad-hoc, inadequate ways, which can have dramatic consequences.”

There are at least three aspects of testing that are interesting for us. First, due to platforms becoming off-the-shelf products, downstream system development activities such as configuration and testing play a central role in platformization processes. Testing in such settings is not primarily for fixing software bugs, as noted by Sommerville: “Systems rarely failed in the sense that they crashed or produced clearly incorrect output. Rather, the failures could only be detected by users who understood their local processes and who could identify where system support was inadequate” (Sommerville, 2008, p. 7). In this way, testing can be seen increasingly as a source of new requirements for the next version of the platforms, helping organizations and vendors cross the “implementation line” (Leonardi, 2009). Second, testing is a collaborative organizational practice. Although software testing has traditionally been regarded as a purely technical activity performed mainly by programmers, several researchers emphasize the importance of involving other stakeholders such as marketing and end-users

(Kawalek and Wood-Harper, 2002; Mäntylä et al., 2012; Rooksby et al., 2009). Idrus et al. (2019) argue that software testing heavily relies on business process and system knowledge as a “vast amount of knowledge is captured, used, shared, stored and reproduced throughout the testing activities.” Ahonen et al. (2004) show how organizational models of the test organization impact the processes and outcomes of testing. Third, software testing is sociotechnical, and artifact-driven as the subject of testing is often the artifacts – here, boundary objects – that result from earlier system design and development activities (Doolin and McLeod, 2012).

Boundary objects in system development

The concept of boundary object (BO) was first introduced by Star and Griesemer (1989), and has since been used in a wide variety of research areas (see e.g., Lee, 2007). Extant literature has emphasized BOs’ importance in collaborating across boundaries and distances – such as organizational, social, and/or cultural – between stakeholder groups or communities (Islind et al., 2019). In the book *Sorting Things Out: Classification and Its Consequences*, Bowker and Star (2000) developed the concept further. Boundary objects are “those objects that both inhabit several communities of practice and satisfy the informational requirements of each of them” (Bowker and Star, 2000, p. 16). Boundary objects are thus “both plastic enough to adapt to local needs and constraints of the several parties employing them, yet robust enough to maintain a common identity across sites” (Star and Griesemer, 1989, p. 393).

In the development of Information systems (IS), BOs are the artefacts that stakeholders interact with in order to facilitate understanding and cooperation across diverse knowledge domains (Doolin and McLeod, 2012). BOs can be technological artifacts such as product versions. They can also be requirement specifications, project timelines, project management methodologies, workshops, drawings, or paper prototypes (Doolin and McLeod, 2012; Islind et al., 2019; Peer and DiSalvo, 2019).

As highlighted by Levina and Vaast (2005), artifacts only become boundary objects when in use by temporarily creating a bridge between groups of different perspectives. In traditional in-house design and development projects, BOs are often used for mutual learning, seeking to overcome the knowledge barriers between those with the technical knowledge and those with domain specific knowledge (Barrett and Oborn, 2010). Islind et al. (2019) suggest that different stages of a platform design process call for different types of BOs to move the collaboration and co-design process forward. Through the different phases the objects move from “open” to “closed”, or unstructured to structured. Open BOs are typically used during the early phase of a design process to give rise to questions and create dialog, while structured objects are introduced later in the process as they are not as easy to change and closer to the final solution.

In platformization processes where a platform is being implemented, the core of the platform is already developed. Boundary objects that represent this core might

therefore seem not as malleable as those in the early stages of system development. Moreover, platformization processes often involve multiple communities of practice – e.g., programmers, healthcare personnel, managers – in large numbers, leading to highly regulated platformization processes. At the same time, while the analysis by Isind et al. (2019) – and many others who have studied BOs in system design – describes how BOs evolve in a phased model of system development, modern iterative methods such as agile (Hoda et al., 2018) are based on the assumption that there is no strict “implementation line” (Leonardi, 2009) dividing design and use. This means that the “closeness” and “openness” of BOs might depend on other factors than their current location in the design process. These factors need to be understood, and we believe employing a practice lens can help us understand and recognize the potential new roles that BOs play in platformization processes.

Case and Method

Case description

In March 2019, the municipality of Trondheim (primary healthcare) and the central Norway health authorities (specialist healthcare) together purchased an Electronic Health Record (EHR) system from the American vendor Epic. This was the culmination of a lengthy process that started ten years earlier with the introduction of the so-called collaboration reform in the Norwegian healthcare (Helse- og omsorgsdepartementet, 2009). Norwegian healthcare is characterized by a strong primary healthcare sector, consisting largely of municipal health services and privately operating general practitioners. The collaboration reform aimed at strengthening primary healthcare even further, moving several duties from hospitals to municipalities. This led to Norway now having one of the most distributed healthcare services in the world (Nylenna, 2020). Due to this distribution – and to prepare for the collaboration reform – the government launched the vision of “One citizen – One medical record” in 2012 (Helse- og omsorgsdepartementet, 2012). By implementing a customized version of Epic’s EHR product, locally called Helseplattformen (Healthcare platform, or HP in this paper), the hope is that the almost 40 000 employees in the involved organizations will share healthcare data, in this way improving healthcare services for the citizens.

Immediately after the purchase of the new EHR product, a new organization was also set up (Figure 1). A new local implementation company called Helseplattformen AS – the middle box in Figure 1 – was created, owned by the hospital and the municipality (new municipalities and other primary care actors are also joining the project and taking part-ownership in this company as they join). HP AS has hired tens of Application Analysts (AAs). AAs are often previous healthcare workers who are in charge of the customization process of the EHR product. In addition, the user organizations – e.g., the hospital and the municipality

–have each set up their local implementation projects that are in charge of organizing the local implementation tasks.

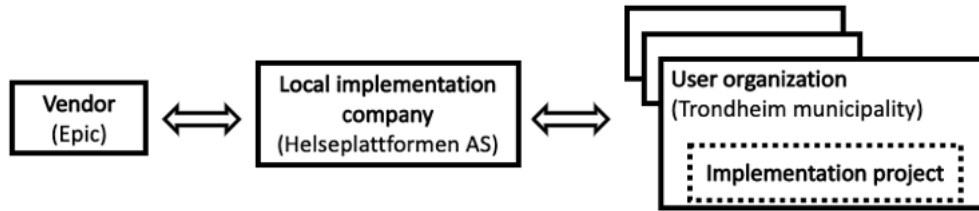


Figure 1. The ecosystem of organizations involved in the digitalization project.

The implementation process for the EHR product consists of six phases as shown in Figure 2. During the specification and development phases, a new version of the EHR product is developed and the assumption is that this new version will fulfill the needs of all the organizations involved. The test and approval phase will test this customized version, and verify its functionality. The study reported in this paper is mainly about the test and approval phase. At the time of writing, all testing is concluded and the training phase has started. The “go-live” date for the product is set for May 1st of 2022.

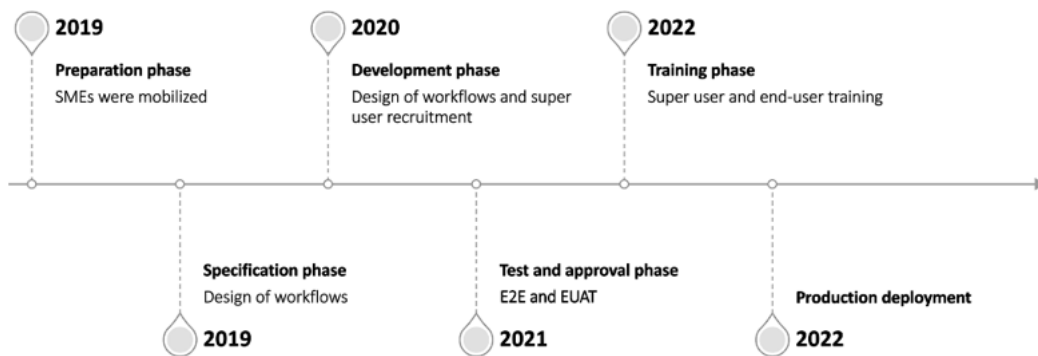


Figure 2. Timeline for the implementation of Helseplattformen.

The above organization and the process of implementing the EHR illustrate a large-scale platformization process as discussed earlier. Early on, it was decided that user participation would be an essential part of the process. This was to avoid past mistakes in e.g., Denmark, where the same product led to a high level of dissatisfaction among healthcare workers after it was implemented. Two types of users have been particularly important for this participation; Subject Matter Experts (SMEs) and Super Users (SUs). Early in the project, the role of SMEs was created to represent domain experts. They have been involved from the preparation phase. More than 400 SMEs, employed in primary or secondary healthcare, have been directly involved in the platform customization during the specification phase. Therefore, it is important to remind the reader that several activities were

conducted with SMEs before the testing phase and that these are not included in the study. In addition, approximately 700 SUs have been recruited to play a key role in training the employees to use the new system.

As discussed previously, testing is a sociotechnical process involving many people in various roles. Three types of tests are in focus in HP (see Table I), End-to-End (E2E) integration tests pass 1 and pass 2, and End-User Acceptance Tests (EUAT). SMEs and SUs have participated in pass 2 of E2E and EUAT. Before the E2E tests, Helseplattformen arranged two separate kick-offs; one for the SMEs and one for the SUs. The aim was to clarify expectations and make them ready to participate in E2E. Trondheim Kommune also arranged a kick-off for the municipal test participants to make sure that they understand their role on behalf of the primary healthcare service. During the kick-off, they were presented with information about how to provide change requests and report on errors. It was also specified that E2E is not training, a demonstration of the solution, or a dedicated test for all possible variations within each integration. The roles of the SMEs and SUs were also explained. The role of the SMEs during E2E is to verify that the adopted solution works, comment on errors and omissions, and reflect on what changes the solution brings to their organization. The SUs are invited, as observers, for two reasons: to build competence before EUAT and to identify changes caused by the solution in their organization. Trondheim municipality also wants to use SUs to identify areas for improvement.

Table I. Overview of the testing activities that involve users in Helseplattformen.

Test	Purpose	Roles
E2E Pass 1	Make sure that the integrations between applications work before involving the users in order to avoid confusion.	Application analyst (AA), Vendor (V)
E2E Pass 2	Verify the regional solution as agreed upon through requirements and adoption, detect errors so that these can be corrected, and identify change requests and areas of improvement.	Application analyst (AA), Subject matter expert (SME), super user (SU), Vendor (V)
EUAT	The SMEs perform the test to verify that key integrations interfaces are working as expected.	Application analyst (AA), Subject matter expert (SME), super user (SU), Vendor (V)

All tests happen in so-called test sessions organized around a test scenario. A test scenario tells a patient's story and involves several applications, workflows, and relevant integrations. For every scenario, there is a test script that describes how the

procedure is performed. Application analysts from Helseplattformen prepare the test scripts in collaboration with SMEs before the testing. Application analysts are in charge of running and facilitating the test sessions, while representatives from the vendor are present in the sessions to answer questions. Figure 3 shows the different testing activities and the people that are involved in each of them. The total number of participants in each test session varies depending on the number of workflows and integrations involved in the scenario. Due to the pandemic, all the test sessions that we observed were organized and run in Microsoft Teams. After each session, the staff from the implementation project in the municipality would hold debriefing meetings with the SMEs and SUs from the municipality.

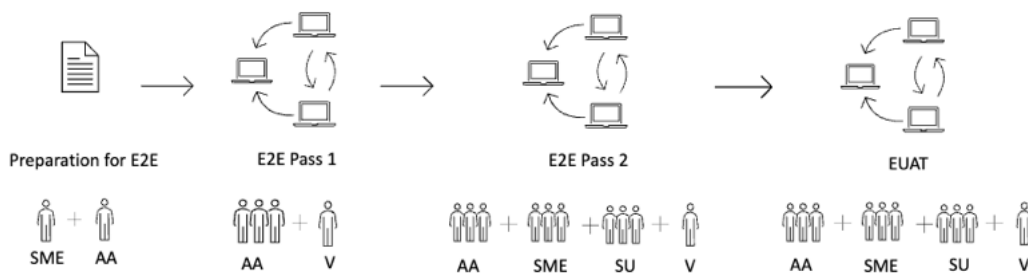


Figure 3. The process of the testing activities in Helseplattformen (see also Table I).

Research method

The strategy to answer the research questions is through an exploratory case study. The research is based on qualitative research methods as the aim is to get an in-depth understanding of boundary objects in the testing phase. Case studies can vary in their approach to time, although this is a short-term study where data was collected over a period of 4 months.

Data was collected through interviews, field observations, and analysis of documents related to the local implementation project in Trondheim municipality. Semi-structured interviews were conducted with stakeholders, including SMEs, SUs, application analysts, and vendor employees. Due to the ongoing pandemic, they were conducted digitally through Microsoft Teams and were, therefore, video recorded by consent. The recordings were then transcribed. Observations were also performed digitally and were documented using extensive field notes. Observations were not recorded.

A qualitative analysis was performed on the data using the qualitative data analysis software, NVivo. The analysis began by identifying codes in the transcriptions, field notes, and documents that were later divided into conceptual categories to identify patterns (Tjora, 2018). These patterns were then used to determine overall themes.

Table II. Data collection methods.

Method	Data
Observation	3 E2E test sessions (8.5 hours in total) 2 Debriefs hosted by the municipality (average of 20 minutes)
Interview	2 SMEs (primary healthcare) 3 Super users (primary healthcare) 2 Application analysts (Helseplattformen AS) 3 Vendor representatives (Epic)
Document analysis	Municipality organizational documentation, public documents, Helseplattformen project documentation, project-related emails, news paper articles

Findings

This section first demonstrates the roles that SMEs and SUs played in testing and what types of breakdowns we observed in their participation in test sessions. We will then present our findings concerning the test version of the products that we observed during the tests and how they supported collaboration.

The role of the end-users in E2E

The goal of E2E testing is to verify the solution as agreed through requirements and adoption. In the SMEs' view, it is difficult to see why they are involved in E2E as they feel it is too late to say something now. However, the vendor gives three reasons why end-users are involved in E2E testing.

They can say, "I don't think that's right" they can say, "oh, that's confusing, but it's the right thing to do, so I need to make sure that gets into training materials," or "I need to tell my colleague that this is going to be what that is." So that is what we expect from the [SMEs]. (Vendor 1)

According to the vendor, E2E testing is not an activity that end-users are usually involved in, but something they decided to do in this project as a preparation for EUAT.

The idea is that if end-users are participating in E2E, their comments [and] questions will be addressed earlier than if we waited for End User Acceptance Testing. This is a strategy we have seen before on other installs. (Vendor 1)

Super users are invited to E2E as observers. Involving SUs in E2E is something that the different health authorities and primary healthcare are invited to do. Still, it

is up to them to decide how they want to use this participation in further work within the health authority or the unit. When asked about their role in the test session, one of the super users responded that

It was to be an observer, so I paid attention to where they [the application analysts] clicked, what they talked about, and then I was asked at the end what I thought about it - how things worked. (Super user 2)

Preparation for E2E

In preparation for E2E testing, SMEs and super users were presented with information about the process and their role through kick-off meetings. All the participants needed to be informed about their role to ensure that testing did not turn into training.

I have seen this across all the implementations I have worked with, end-users would love to be on the phone the entire time of the implementation, but they also have their own jobs to do. We want to make sure they are still providing patient care, and there is going to be a whole other effort devoted to them getting their training, so making sure that we are using testing to test the system. (Vendor 1)

One of the SUs we interviewed had participated in the kick-off meetings intended to introduce them to the process of E2E. However, it seemed like the information they received led to more confusion than clarity for some.

During the kick-off, I got the impression - there was kind of two messages [about my role] - one was that I was supposed to be an observer and the other one was that we were supposed to have something to write with and sort of take notes along the way. (Super user 1)

An SME also mentioned that the information was presented in a terminology that they were not familiar with:

They speak in a language that I don't understand at all with English computer expressions that are completely foreign to me. It might not even be computer expressions, but it's a completely foreign terminology. (SME 2)

The SME, therefore, spent a lot of time in advance to prepare for the test by trying to go through the script several times and look at the workflows that were involved. The foreign terminology used in the kick-off also made it seem like the threshold for making comments during the test was high.

The threshold for notifying change seemed to be quite high, and they [the test team] talk about systems and ways to notify changes that I have never heard of before, so it gets kind of like s***, how bad does it have to be before one dares to say something? (SME 2)

According to an SME, there was a long sequence at the beginning of the test session where none of the SMEs or SUs dared to say anything, and everyone was "holding their breath" before it gradually became easier to take the floor and speak up.

Breakdowns in the process

The Helseplattformen (HP) project has two process levels: the process introduced by the vendor and the local implementation process (see also Figure 1). Sometimes these are difficult to coordinate. While HP AS arranges the E2E tests, the municipality is invited to involve super users, which leads to these process levels intertwining. The test session did not clear up the uncertainties created by the information received by the end-users in advance. One of the SUs mentioned that the super users were not introduced at the beginning of the test session in the same way as the other participants, which led to further confusion about why they were there.

There was just a bit of uncertainty about my role and what I was really doing there. Because I sat and took notes, and there was a lot, but then again that might be about content, and that might not be what they were supposed to be testing at that moment. But if that is not what they were testing, there was kind of no need for me to participate. (Super user 1)

After each test session, the SMEs and SUs are invited to a debrief hosted by Trondheim municipality to share their immediate thoughts and experiences. Both SMEs and SUs are also asked to fill in a questionnaire to help them identify changes in their day-to-day work routine using the new solution. Although they express that they find the debrief useful, they are unsure whether the input will be taken further by HP AS and lead to changes. Still, they say that it was nice to be able to share their experiences right away.

During one of the test sessions, an SME discovered that a word that had been translated from English to Norwegian had not been translated to the word that the SME wanted it to be. The SME decided not to say anything about it during the test because it did not have anything to do with the integration with other applications. The questionnaire had no field to report about the translation either, which made the SME unsure of where to provide this feedback. The SME decided to e-mail the test team in Trondheim municipality to ask who should receive such feedback. In the e-mail, the SME stated that there was a need to provide some feedback regarding word choices and some other small changes in the solution and that the SME did not remember how to do it. The SME ended the e-mail by asking, "should I send it to

the application analyst? If so, which one?". After receiving an insufficient response from the municipality, the SME decided to contact one of the application analysts directly. The application analyst thanked the SME for not bringing the questions up during the test and cleared up the SME's uncertainties.

I could have asked about it during the test, but it didn't have anything to do with the integration between the different applications, so that's why I couldn't bring it up. (SME 1)

According to the SME, a change for that translation was requested a year ago, but the discovery makes the SME uncertain whether it has been implemented since the test version of the product might not include all change requests.

The product in E2E testing

E2E testing in Helseplattformen involves different BOs like the test scripts, test plans, and the version of the product - where the latter seems to be the most important. The E2E test session is the first time the super users see a version of Helseplattformen live. Despite the uncertainty related to their role, all super users reported a growing curiosity and increased motivation to use the new solution after seeing the test version of the product. The super users also thought seeing the application analysts reporting errors so thoroughly was reassuring.

Now I have seen an example from Helseplattformen, and it looks like we will get a lot of information about a patient on one page. The way the journal system works today, you have to go into separate journal codes to find the information you need. Nothing is gathered in one place. So I am looking forward to getting lots of information gathered in one place. (Super user 2)

Since the scenario that was tested in one of the sessions included both the hospital side and municipal health services, one of the super users mentioned that seeing the product provided them with insight into the information flow between the hospital and the unit.

For me, it was more insight into what is going on in secondary healthcare, but also the information that is sent from there [the hospital] to the municipality, via the Health and Welfare Office and to the Health Center. So you get some insight into a process that you don't know much about. You know that it is done, but not always what kind of information is sent. Except for a discharge letter or a final note. (Super user 3)

After seeing the product another super user became more curious but states that it may not have been necessary for the super users to attend an entire test session to get the same effect.

It made me a little curious about what it's like to use it. I got a quick glimpse of the screen, so it's a bit exciting, but I think I would have thought the same even if for example, the leading super user in my department had shown us that we can do this and this [...], and this is what it looks like for us, I think that would have been the same as the test. (Super user 1)

The test is also described as a "wake-up call" by one of the super users saying that

I believe everyone would benefit from participating in a test like this because [...] the process has lasted for so long, and it is so peripheral to the employees. [...] I am afraid that people forget that this is actually coming. So it is sort of a wake-up call that this is happening and that we just have to keep up. (Super user 2)

Communication through the product

According to a vendor representative, the challenge of communication in a large project appears in the testing phase the most as that is where different pieces of the software and different people are working together really for the first time. An application analyst gives examples of typical comments that emerge during test sessions and says that

It can be 'why did you press this button and not that one' and the answer is often that it doesn't matter because you are taken to the same [place]. [...] Or it can be that a flow chart is not exactly how it was agreed upon, which might be because it has not been fully developed, but it has been built just enough for it to be tested, so questions, and comments like that may emerge. (Application analyst 1)

During a test, one of the SMEs discovered that there was a person in a patient list that was too young to be in that specific list and asked why that was the case. "It is probably just fill [test data]," the application analyst replied, "but you can press the remove patient button if that is the case." Some of the other end-users comment "good observation" in the chat, to which the SME reply, "well, this is our future, and it is a very realistic situation".

The super user that was taking notes also asked to share a few comments as the end of the test session was approaching, to which the Helseplattformen test team responded, "bring them on". The super user refers to a specific part of the solution and asks whether it is okay to write notes to themselves that are confidential there. An application analyst replies, "yes, but nothing that can't stand the light of day". The super user then asks whether it is okay to write the patient's name or if you have to use a number or ID. "No, you can write their name" the application analyst replies and mention that Helseplattformen is a secure system. One of the test team members proceeded to say that it is common to get a feeling of uncertainty when

seeing the product like this for the first time and that there will be a debrief for the municipality shortly after the test session is finished. The super user responds by saying, "then I will save my comments until then".

Discussion

Our findings demonstrate some well-known challenges related to design artifacts as boundary objects (Doolin and McLeod, 2012; Islind et al., 2019). For instance, we see that BOs, particularly the test product itself, are efficient communication tools. The product engages the users because it communicates to them. Consistent with Islind et al. (2019), the product is a closed BO and triggers more detailed conversations that would likely not have emerged with less concrete BOs. At the same time, due to the fact that our case is about the testing phase of a customized platform, we also see some differences compared to the development of new platforms and systems. These differences relate to both the boundary object itself - here, the test product - and the collaboration processes based on that object.

The role of boundary objects in platform testing

The super users saw the product for the first time in the test sessions. Even though they were confused about what they could comment on and when, all of them reported a growing curiosity and motivation to use the end product. At the same time, this "first meeting" was the source of the confusion; can they comment on the product's obvious faults that shine through? If yes, how? This situation is illustrative of the discontinuity involved in designing and using large platform products. The BO - the product - is closed, but it was not closed by these users. It was closed mainly by users in earlier projects in USA and the SMEs and developers in earlier phases of this project. While the product as a BO invites modifications - because it is so material and easy to understand and scrutinize by the super users - it is also "protected" by contracts and change management regimes in the project. SMEs who participated in the earlier phases of the project report similar issues with what they could comment on (Zahlsen et al., 2020).

Some of the confusion we observed was related to the fact that in large development projects, the product is continuously being developed and modified - i.e., the BO might be perceived as closed by end-users but is still open for others in the project. The end-users often do not know how mature or complete a product is (a point also raised by Dourish et al., 2020). This might mean that a change that is seen as important by a super user is regarded as negligible by an application analysis or the vendor. Moreover, the platform is a special type of layered BO because it consists of parts that can be changed easily, parts that will cost some resources to change, and parts that cannot practically and financially be changed in one project. However, for users, it is difficult to know which is which. This layered nature leads to a knowledge management problem. Often the distance between those who know the different layers of the BO is so large that it is practically not

feasible to have all of them in the same room. Local discussions can therefore become confused and error-prone as those participating do not know the BO well enough.

Our findings suggest that the phased “closing” process of BOs, as described by Islind et al. (2019), might not be suitable for large-scale platformization projects. This closing process is often punctuated in platformization processes by the passing of time and the involvement of vastly different users in the various phasing of the closing process (Pollock et al., 2007). BOs in use are often understood as “technical” by different users, meaning that they are built using the language of another community of practice. In our case, this was exacerbated by the fact that we are dealing with an American product being adapted to a Scandinavian setting. While in smaller projects developing new platforms, such “technicalities” can be sorted out by the group, in large and longitudinal projects, they often do not get solved immediately and carry on creating confusion in the cooperation. This is partly because the BO can only be modified through a protective shield of “change management” tools and processes, making the threshold to resolve inconsistencies very high. Even if some local understanding is created by some people using ad hoc means, the longevity of the project means that those people are replaced by new ones in the next phase, and the original or new misunderstandings appear again.

Traditionally product testing is regarded as the end-point of a long and laborious design and development process. This is the consequence of following a waterfall development process – often mandated by regulations surrounding public procurement. At the same time, the testing phase is often the first time users see the product. While the vendor is interested in “closing” the BO and signing the final contract papers, the users start their journey first during the product testing sessions. When the end-users get involved, the product becomes the main BO to create a shared understanding. As mentioned by a vendor representative, the challenge of communication in a large project appears in the testing phase the most as this is where the different pieces of the software and the different communities of practice are brought together for the first time. In this sense, product testing needs more attention as a practical and central tool to cross the “implementation line” and open up the black box of needed functionality (Leonardi, 2009).

Collaborative processes surrounding the boundary objects

Due to the size of the project and the number of people involved, access to the BO – i.e., the test product – was controlled by strict and formal processes. These were standardized processes introduced into the project by the vendor. Various kick-off meetings and workshops were held to train the SMEs and super users in the processes. Despite the training, these processes and the associated roles were not understood well by those taking on these roles. The introduction of formal processes created confusion about what could be done with the BO and when. For instance, the SMEs often did not know whether they could comment on the BO directly

during the test session or whether it was “too late.” When the user representatives went outside the formal process, e.g., sending emails directly to the test session facilitators, it was not clear what happened to such communication. This, in turn, led to more insecurity about the role of the user representatives in the process.

The ownership of the process by the vendor made it hard for local user organizations to change it to fit their needs. At the same time, valuable knowledge from local processes was not integrated into the formal process and seemed to get lost. Local processes such as the debrief meetings in the municipality often created a more trustful environment for the exchange of ideas. However, we observed numerous breakdowns in the transfer of knowledge from these local processes to the project-wide formal processes. Moreover, although the municipality had implemented additional local processes, these processes were sometimes influenced by the project-wide processes. For instance, the questionnaires that were distributed by the municipality among their SMEs and super users did not open for commenting the BO – even if the debrief meetings demonstrated that there still was a need for it locally. Instead, the questionnaires asked the user representatives to only comment on how they should change their own work processes to accommodate for the new platform. We believe the formal processes in the project could have benefited greatly by taking into account local processes as valuable input to the project.

In large-scale platformization processes, BOs are surrounded by numerous restrictions. Adapting BOs in these projects becomes a costly task. Instead of doing quick changes locally, the project team often needs to propagate the change back to a platform core owner, who needs to calculate the impact of the change for all users of the core and calculate the cost of the change for this customer. Changes are furthermore regulated by formal contracts, which restrict the allowed cost. This means that changes that make sense in a local context are left in a “product backlog” and may not be relevant once the project has moved on and the level of knowledge and maturity has changed, making these changes irrelevant or even harmful. This is far from what happens in small-scale new platform development as reported in e.g., Farshchian and Thomassen (2019) and Islind et al. (2019). Once more, product testing has greater potential than being regarded as the last phase in a project. Regarding testing as a sociotechnical process and the test product as a BO early on can provide a valuable vehicle not only for requirements but also for training and preparedness.

Conclusion

Boundary objects play a central role in testing as educational and motivational objects as well as “wake-up calls” for new users. The product as a BO creates enthusiasm among the user representatives. As opposed to other more abstract BOs such as plans, test scripts, and presentations, the test product seems to be much easier to relate to own work and plays a central role in creating reactions. Lack of local ownership and strict, formal processes reduce a BO’s spanning properties. It

is easier for user representatives to feel safe in more local meetings involving other municipality employees or preferably their own unit. Useful information about the process and the BO emerge in such meetings, which might mean that the process of testing should have been more distributed and performed locally by local SMEs and super users. Giving user representative earlier access to the product by starting product testing earlier, and establishing local ownership of the testing processes, can help reduce various risks that often face large-scale platformization projects.

Our data collection and analysis should be considered as ongoing. This paper presents the results of a first round of thematic analysis, and an initial theoretical framing. The hope is therefore that this paper will provide us the opportunity to present and receive feedback on this work in progress.

Acknowledgements

We thank employees from the municipality of Trondheim, Helseplattformen AS, and Epic Systems Corporation for providing us the opportunity to observe and interview them, and for their feedback on earlier versions of this paper. We also thank the members of Forskerfabrikken at NTNU and the anonymous reviewers of ECSCW for their comments, which helped improved this paper. The research reported here was partly funded by the PlatVel project (www.ntnu.edu/platvel).

References

- Abraham, R. (2013): 'Enterprise architecture artifacts as boundary objects-A framework of properties'. In: *21st European Conference on Information Systems, ECIS 2013*.
- Ackerman, M. S., J. Dachtera, V. Pipek, and V. Wulf (2013): 'Sharing knowledge and expertise: The CSCW view of knowledge management'. *Computer Supported Cooperative Work (CSCW)*, vol. 22, no. 4-6, pp. 531–573.
- Ahonen, J. J., T. Junttila, and M. Sakkinen (2004): 'Impacts of the organizational model on testing: Three industrial cases'. *Empirical Software Engineering*, vol. 9, no. 4, pp. 275–296.
- Ajzen, M. and L. Taskin (2021): 'The re-regulation of working communities and relationships in the context of flexwork: A spacing identity approach'. *Information and Organization*, vol. 31, no. 4, pp. 100364.
- Akkerman, S. F. and A. Bakker (2011): 'Boundary crossing and boundary objects'. *Review of educational research*, vol. 81, no. 2, pp. 132–169.
- Barrett, M. and E. Oborn (2010): 'Boundary object use in cross-cultural software development teams'. *Human Relations*, vol. 63, no. 8, pp. 1199–1221.
- Bertolino, A. (2007): 'Software testing research: Achievements, challenges, dreams'. In: *Future of Software Engineering (FOSE'07)*. pp. 85–103.
- Bowers, J. M. (1991): 'The Janus Faces of Design: Some Critical Questions for CSCW'. In: J. M. Bowers and S. D. Benford (eds.): *Studies in Computer Supported Cooperative Work: Theory, Practice and Design*. Amsterdam, etc., pp. 333–350, North-Holland.

- Bowker, G. C. and S. L. Star (2000): *Sorting things out: Classification and its consequences*. MIT press.
- Bygstad, B. and O. Hanseth (2019): *Transforming Digital Infrastructures through Platformization*. Association for Information Systems.
- Cordella, A. and A. Paletti (2019): ‘Government as a platform, orchestration, and public value creation: The Italian case’. *Government Information Quarterly*, vol. 36, no. 4, pp. 101409.
- Dahl-Jørgensen, T. C. and E. Parmiggiani (2020): ‘Platformization of the public sector: Assessing the space of possibility for participation’. In: *PDC’20: Participatory Design Conference 2020-Participation Otherwise*.
- De Reuver, M., C. Sørensen, and R. C. Basole (2018): ‘The digital platform: a research agenda’. *Journal of Information Technology*, vol. 33, no. 2, pp. 124–135.
- de Reuver, M., C. Sørensen, and R. C. Basole (2018): ‘The Digital Platform: A Research Agenda’. *Journal of information technology*, vol. 33, no. 2, pp. 124–135.
- Doolin, B. and L. McLeod (2012): ‘Sociomateriality and boundary objects in information systems development’. *European Journal of Information Systems*, vol. 21, no. 5, pp. 570–586.
- Dourish, P., C. Lawrence, T. W. Leong, and G. Wadley (2020): ‘On being iterated: The affective demands of design participation’. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. pp. 1–11.
- Drewel, M., L. Özcan, C. Koldewey, and J. Gausemeier (2021): ‘Pattern-based development of digital platforms’. *Creativity and Innovation Management*, vol. 30, no. 2, pp. 412–430.
- Ellingsen, G. and M. Hertzum (2019): ‘User participation in the implementation of large-scale suite systems in healthcare’. In: *Infrahealth 2019-Proceedings of the 7th International Conference on Infrastructures in Healthcare 2019*.
- Faraj, S. and S. Pachidi (2021): ‘Beyond Uberization: The co-constitution of technology and organizing’. *Organization Theory*, vol. 2, no. 1, pp. 2631787721995205.
- Farshchian, B. A., M. Grisot, S. Hochwarter, A. Sigríður Islind, M. Mikalsen, E. Parmiggiani, and P. Vassilakopoulou (2021): ‘Platform-Supported Cooperative Work’. In: *Proceedings of 19th European Conference on Computer-Supported Cooperative Work*.
- Farshchian, B. A. and H. E. Thomassen (2019): ‘Co-creating platform governance models using boundary resources: A case study from dementia care services’. *Computer Supported Cooperative Work (CSCW)*, vol. 28, no. 3, pp. 549–589.
- Gawer, A. (2014): ‘Bridging differing perspectives on technological platforms: Toward an integrative framework’. *Research policy*, vol. 43, no. 7, pp. 1239–1249.
- Gawer, A. (2022): ‘Digital platforms and ecosystems: remarks on the dominant organizational forms of the digital age’. *Innovation*, vol. 24, no. 1, pp. 110–124.
- Gerson, E. M. and S. L. Star (1986): ‘Analyzing due process in the workplace’. *ACM Transactions on Office Information Systems*, vol. 4, no. 3, pp. 257–270.
- Ghazawneh, A. and O. Henfridsson (2013): ‘Balancing platform control and external contribution in third-party development: the boundary resources model’. *Information systems journal*, vol. 23, no. 2, pp. 173–192.

- Helse- og omsorgsdepartementet (2009): 'Samhandlingsreformen: rett behandling - på rett sted - til rett tid'. <https://www.regjeringen.no/no/dokumenter/stmeld-nr-47-2008-2009-/id567201/>.
- Helse- og omsorgsdepartementet (2012): 'Én innbygger – én journal'. <https://www.regjeringen.no/no/dokumenter/meld-st-9-20122013/id708609/>.
- Helseplattformen (2019a): 'Bakgrunn for Helseplattformen'. <https://helseplattformen.no/bakgrunn-for-helseplattformen>.
- Helseplattformen (2019b): 'Prosjektet'. <https://helseplattformen.no/om-oss/prosjektet>.
- Hoda, R., N. Salleh, and J. Grundy (2018): 'The rise and evolution of agile software development'. *IEEE software*, vol. 35, no. 5, pp. 58–63.
- Huvila, I., T. D. Anderson, E. H. Jansen, P. McKenzie, L. Westbrook, and A. Worrall (2014): 'Boundary objects in information science research: An approach for explicating connections between collections, cultures and communities'. *Proceedings of the American Society for Information Science and Technology*, vol. 51, no. 1, pp. 1–4.
- Idrus, H. M. et al. (2019): 'Tacit Knowledge in Software Testing: A Systematic Review'. In: *2019 6th International Conference on Research and Innovation in Information Systems (ICRIIS)*. pp. 1–6.
- Islind, A. S., T. Lindroth, J. Lundin, and G. Steineck (2019): 'Co-designing a digital platform with boundary objects: bringing together heterogeneous users in healthcare'. *Health and Technology*, vol. 9, no. 4, pp. 425–438.
- Jarrahi, M. H., G. Newlands, M. K. Lee, C. T. Wolf, E. Kinder, and W. Sutherland (2021): 'Algorithmic management in a work context'. *Big Data & Society*, vol. 8, no. 2, pp. 20539517211020332.
- Johansen, R. (1988): *Groupware. Computer Support for Business Teams*. New York and London: The Free Press.
- Kawalek, P. and T. Wood-Harper (2002): 'The finding of thorns: user participation in enterprise system implementation'. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, vol. 33, no. 1, pp. 13–22.
- Krabbel, A. and I. Wetzel (1998): 'The customization process for organizational package information systems: A challenge for participatory design'. In: *Proceedings of the PDC*, Vol. 98. pp. 45–54.
- Lee, C. P. (2007): 'Boundary negotiating artifacts: Unbinding the routine of boundary objects and embracing chaos in collaborative work'. *Computer Supported Cooperative Work (CSCW)*, vol. 16, no. 3, pp. 307–339.
- Leonardi, P. M. (2009): 'Crossing the implementation line: The mutual constitution of technology and organizing across development and use activities'. *Communication Theory*, vol. 19, no. 3, pp. 278–310.
- Levina, N. and E. Vaast (2005): 'The emergence of boundary spanning competence in practice: Implications for implementation and use of information systems'. *MIS quarterly*, pp. 335–363.
- Mäntylä, M. V., J. Itkonen, and J. Iivonen (2012): 'Who tested my software? Testing as an organizationally cross-cutting activity'. *Software Quality Journal*, vol. 20, no. 1, pp. 145–172.

- Myers, G. J., C. Sandler, and T. Badgett (2012): *The art of software testing*. Hoboken, N.J: John Wiley & Sons, 3rd ed edition.
- Nicolini, D., J. Mengis, and J. Swan (2012): 'Understanding the role of objects in cross-disciplinary collaboration'. *Organization science*, vol. 23, no. 3, pp. 612–629.
- Nylenna, M. (2020): *Helsetjenesten i Norge: et overblikk*. Oslo: Gyldendal, 2. utgave. edition.
- Oates, B. J. (2006): *Researching information systems and computing*. London: Sage Publications.
- Orso, A. and G. Rothermel (2014): 'Software Testing: A Research Travelogue (2000–2014)'. In: *FOSE 2014*. New York, NY, USA, p. 117–132, Association for Computing Machinery.
- Peer, F. and C. DiSalvo (2019): 'Workshops as Boundary Objects for Data Infrastructure Literacy and Design'. In: *Proceedings of the 2019 on Designing Interactive Systems Conference*. pp. 1363–1375.
- Plantin, J.-C., C. Lagoze, P. N. Edwards, and C. Sandvig (2018): 'Infrastructure studies meet platform studies in the age of Google and Facebook'. *New media & society*, vol. 20, no. 1, pp. 293–310.
- Poell, T., D. Nieborg, and J. Van Dijck (2019): 'Platformisation'. *Internet Policy Review*, vol. 8, no. 4, pp. 1–13.
- Pollock, N., R. Williams, and L. D'Adderio (2007): 'Global software and its provenance: generification work in the production of organizational software packages'. *Social studies of science*, vol. 37, no. 2, pp. 254–280.
- Rodon Modol, J. and B. Eaton (2021): 'Digital infrastructure evolution as generative entrenchment: The formation of a core–periphery structure'. *Journal of Information Technology*, vol. 36, no. 4, pp. 342–364.
- Rooksby, J., M. Rouncefield, and I. Sommerville (2009): 'Testing in the wild: The social and organisational dimensions of real world practice'. *Computer Supported Cooperative Work (CSCW)*, vol. 18, no. 5-6, pp. 559.
- Schrieck, M., M. Wiesche, and H. Krcmar (2021): 'Capabilities for value co-creation and value capture in emergent platform ecosystems: A longitudinal case study of SAP's cloud platform'. *Journal of Information Technology*, vol. 36, no. 4, pp. 365–390.
- Sommerville, I. (2008): 'Construction by configuration: Challenges for software engineering research and practice'. In: *19th Australian Conference on Software Engineering (ASWEC 2008)*. pp. 3–12.
- Sommerville, I. (2016): *Software engineering*. Pearson, 10th ed., global ed. edition.
- Star, S. L. and J. R. Griesemer (1989): 'Institutional ecology, translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39'. *Social studies of science*, vol. 19, no. 3, pp. 387–420.
- Strong, D. M. and O. Volkoff (2010): 'Understanding Organization—Enterprise system fit: A path to theorizing the information technology artifact'. *MIS quarterly*, pp. 731–756.
- Thomas, L. D. and R. Tee (2021): 'Generativity: A systematic review and conceptual framework'. *International Journal of Management Reviews*.

- Tiwana, A. (2014): *Platform Ecosystems: Aligning Architecture, Governance, and Strategy*. Morgan Kaufmann.
- Tjora, A. (2018): *Qualitative research as stepwise-deductive induction*. Routledge.
- Trondheim kommune (2021): 'Helseplattformen'. <https://www.trondheim.kommune.no/helseplattformen/>.
- Vassilakopoulou, P., M. Grisot, T. Blegind Jensen, N. Sellberg, J. Eltes, A. A. Thorseng, and M. Aanestad (2017): 'Building national ehealth platforms: the challenge of inclusiveness'. In: *Proceedings of the International Conference on Information Systems-Transforming Society with Digital Innovation, ICIS 2017*.
- Williams, R. and N. Pollock (2012): 'Research commentary—moving beyond the single site implementation study: how (and why) we should study the biography of packaged enterprise solutions'. *Information Systems Research*, vol. 23, no. 1, pp. 1–22.
- Zahlsen, Ø. K., D. Svanæs, A. Faxvaag, and Y. Dahl (2020): 'Understanding the Impact of Boundary Conditions on Participatory Activities.'. In: *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*. pp. 1–11.
- Øverby, H. (2018): 'Digital economics : how Information and communication technology is shaping markets, businesses, and Innovation'.

Corentin Stalder, Myriam Lewkowicz, Guillaume Ducellier (2022): The potential of a digital twin to design computational coordination mechanisms. The case of the French railway infrastructure. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_ep08

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_ep08

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

The potential of a digital twin to design computational coordination mechanisms. The case of the French railway infrastructure

Corentin Stalder, Myriam Lewkowicz, Guillaume Ducellier

SNCF Réseau, Troyes University of technology (UTT), LIST3N/Tech-CICO

Contact Author: corentin.stalder@utt.fr myriam.lewkowicz@utt.fr; guillaume.ducellier@utt.fr

Abstract.

The digital twin is a key technology of Industry 4.0. It offers a digital replica of a physical system. As part of research work with SNCF Réseau, we have sought to explore the role the digital twin can play to support coordination mechanisms. We conducted a multi-sited ethnography, which allowed us to define a technology probe to explore our research question in-situ. This exploratory work shows that the digital twin enables the rapid design of computational coordination mechanisms. Furthermore, we raise the point that a digital twin can bind descriptive information about the physical system and information used in coordination practices to move towards effortless coordination.

Keywords: Digital twin, railway infrastructure, coordination mechanisms, technology probe.

1 Introduction

The digital twin has been a trending topic in academic research and the industry since 2017. More than 2500 papers¹ were published on this topic in 2021. Gartner identified the digital twin as the fourth most strategic technology trend in 2018, ahead of virtual reality (Gartner, 2017).

The digital twin is a multi-scale, digital representation of a physical system, updated in real-time and defined as a mirror of the physical system (Tao et al., 2018a).

It has emerged in the product lifecycle management (PLM) community and aims to overcome the lack of digital continuity that may exist between the product design and its management, while taking advantage of the possibilities offered by IoT and big data.

However, even if one of the objectives of the digital twin was to enhance cross-business cooperation (Grieves, 2015), few socio-technical studies have studied it (Tao et al., 2019b; Lamb, 2019). We propose here an exploratory study conducted at SNCF Réseau, where a digital twin of the railway infrastructure has been built (Costa D'Aguiard et al., 2019). Our aim is to explore the potential of the digital twin to support coordination among the different actors involved in the maintenance of the railway infrastructure. Our research falls within the digital twin industrial program of SNCF Réseau that has been leading the railway system's digital twin development for three years now (Gautier et al., 2019).

In order to study the potential of the digital twin to support the coordinative practices in place for managing the infrastructure, we have conducted a multi-sited ethnographic study of the work scheduling activity at SNCF Réseau. We have chosen to focus on this activity due to its criticality and the multiplicity of actors involved. This study allowed us to clarify the organizational context and to identify the coordinative practices implied. We then used the coordination mechanisms (Schmidt and Simone, 1996) framework, as it conceptualizes a socio-material entanglement of articulation work to analyze fieldwork practices and computerized coordination mechanisms (CCM) to support articulation work. From this analysis, we have designed a technology probe (Hutchinson et al., 2003) using the digital twin to prospect its technological capability to support the articulation work that is involved in the work scheduling activity.

2 Industrial Context

SNCF Réseau is the French railway infrastructure manager. Its task is to operate, maintain and develop the 30,000 kilometers of lines that compose the railway network (Ministère de l'économie et des finances, 2019). The railway network is a complex system, first because of its geographic geographical extent, and secondly

¹ The number of articles is calculated from the analysis of the search result on Scopus with the query: TITLE-ABS-KEY ("digital twin")

because of the number of subsystems that compose it (rails, energies, signaling, infrastructure...). This system is ageing and therefore requires a lot of maintenance work. Indeed, the railway network is at the heart of current concerns for low-carbon mobility in Europe and it represents a crucial mobility hub, with highly exploited lines, which has accelerated the degradation process. On top of that, the 2024 Olympics games in Paris will increase mobility needs. Thus, in 2019, 2.7 million euros were invested in the infrastructure, which has led to the renewal of 942 km of lines. In addition, 2.2 million euros were invested in network development projects (electrification, creation of new lines, increase of existing lines' capacity, etc. . .), which has also resulted in infrastructure works (Guinet and De Nicolay, 2019). To face this situation and to improve the management of its infrastructure, the company is engaged in a major transformation of its activity. This transformation is both organizational and digital. One of the digital transformations is the development of a digital twin. SNCF Réseau defined the digital twin of the railway infrastructure as follows (Costa D'Aguiard et al., 2019, p. 3):

[The digital twin is] a digital replica of a physical system, connected to the railway infrastructure and work practices. Its input comes from the reference databases, and sensors fitted in the infrastructure that serve to keep this data up-to-date in real time throughout the system life cycle. This digital replica is designed to simplify the decision-making process, and is structured through three service levels (information, analysis and configuration management). Each rail system user/persona will have access to different views (dashboards, spreadsheets, forms, 2D diagrams, 3D objects, etc.) of the digital twin depending on their functions and access rights.

In other words, the digital twin of the railway system ensures the description of the infrastructure, the recording of all the operations done on the infrastructure, feedback from the field through reports and sensors to update the digital twin and support for decision-making through data analytics. The digital twin of the railway system is composed of the data model, the data warehouses, the update process, and the end-user applications. The data model is now defined and has been implemented in data warehouses. A manual update process is in place and several end-user applications are in production.

The data model aims to be a unified and unique description model of the railway system: a topo-functional model (International Railway Solution, 2016). In other words, the model defines the objects of the infrastructure from a functional point of view and the topological relationships between objects. This model structures data warehouses. Data warehouses are a set of databases exposing data through REST web services. Each object in the model is associated with a web service. Moreover, with the data relationships, it is possible to navigate between objects using objects relationships. For example, it is possible to know tracks composing a line with the line request. It allows the application designer to get a rich description of the infrastructure with different levels of abstraction. It enables the designer to adapt to a wide variety of needs while using a single data source. In the end, users can enjoy a wide variety of representations that are all consistent, and with possible bridges among them. (International Railway Solution, 2016)

Data warehouses are updated with a manual process. It means that the update is done manually; data contained in documents are rewritten in the data warehouses. Finally, the updated data are used in the end-user applications. The end-user applications focus on specific use cases and use data from data warehouses.

3 Related work

In this section, we first present the technological concept of the digital twin, emphasizing the few works that foresee this technology for supporting cooperative work and therefore the originality of the program launched by SNCF Réseau in which we participate.

We then focus on the challenges when engaging with computational coordination artifacts to frame our analysis of the digital twin capability to produce "in-between" coordinative artifacts.

3.1 Digital twin

The digital twin, a digital replica of a physical system, is a key concept of industry 4.0 (Pires et al., 2019). Grieves offered the first conceptualization in 2015, defining a digital twin as a "digital equivalent to a physical product" (Grieves, 2015, p. 1), and structuring it in three main parts:

- The physical products in the real space, for example a rocket or the railway infrastructure.
- The virtual products in the virtual space, for instance a digital mock-up of the product, or a database containing all the data related to the properties, the behavior, and the state of a product.
- The connections between data and information that link the virtual and the real spaces together.

Following Grieves' work, Tao and Zhang (2017) worked on a multi-scale, multi-source and multi-sector representation. On top of the data, the virtual entity and the physical entity, Tao and Zhang introduce a service layer to the digital twin concept. The digital twin can then be described through five dimensions: (1) the data is the center, it drives (2) the physical system (composed of humans and machines), (3) the digital system (the digital model of the system) and (4) the service system (the enterprise information system, computer tools ...). Each of these dimensions communicate through a (5) data flow in order to achieve a continuous optimization of the shop floor (Tao and Zhang, 2017, p. 20420-20421).

In other words, a digital twin refers to a virtual copy (a replica) of a physical system, connected to it via sensors; it permits the fusion of data coming from the sensors put on the physical system and data coming from simulations. Finally, the service layer ensures the management of both the physical and the virtual entities. In the digitization program of SNCF Réseau in which we are involved, the physical system is made of the railway assets (tracks, catenaries ...) and the workers

(maintainers, switchmen...), the virtual system can be a 3D point cloud linked to a digital mock-up and the service system can be the train scheduling system, the maintenance management system, etc.

Even if the development of the digital twin obviously presents technological challenges (how to merge different data sources, how to update the digital twin in real time ...), it also presents socio-technical challenges. In particular, how to build representations that are intelligible for users and how these representations could ease complex cooperative work.

So far, these socio-technical challenges do not seem to be fully addressed in the literature. Indeed, despite the digital twin being built in order to offer a real-time, integrated and collaborative representation of the infrastructure (Tao et al., 2019b; Grieves, 2015), little research work addresses the use of the digital twin as a means to communicate, coordinate or cooperate (Tao et al., 2019b; Lamb, 2019). Of the 471 articles on the digital twin studied by Lamb (2019), 27% were focussed on performance issues (efficiency, productivity ...), 22 % were product-centered (asset health monitoring, life cycle management ...), only 8% were interested in the potential of the digital twin for collaboration and 8% were adopting a human-centered approach.

However, we can see a recent evolution, as shown in the work presented by Nochta et al. (2021)], that offers a socio-technical study of the city of Cambridge's digital twin. This work tackles the difficulty of building representations that are understandable by users of digital twin models.

In our work, we aim to contribute to this discussion regarding the potential of the digital twin by addressing how the railway system's digital twin could offer a common representation allowing more effortless coordination of work conducted by the diverse actors involved in its maintenance. Hence, even if the CSCW research community has studied CAD and BIM systems (Møller and Bansler, 2017; Bjørn et al., 2021; Schmidt and Wagner, 2004), to our knowledge, the potential of the digital twin technology to support cooperation has never been addressed. However, we believe that the digital twin, such as BIM, can offer new opportunities to make cooperative work in complex industrial settings more practical.

SNCF Réseau's digital twin vision is closed to Tao's model. Thus, the digital twin of the railway infrastructure is not a one-piece application, but a data and service network aiming to facilitate infrastructure management and cooperation among services. This digital twin paradigm is close to the development of an information infrastructure (Susan Star, 1999). The coupling between artifacts is automated and systematized, which makes the relationships between artifacts invisible (Bossen and Markussen, 2010).

3.2 Supporting coordination - the potential role of the digital twin

The articulation work is all the work that consists of adapting to unforeseen events (Star and Strauss, 1999); it is a fundamental component of cooperative work. The complexity of articulation work increases with the number of actors, the

distribution of work and the interdependencies between the actors. To deal with this complexity, artifacts are used (Schmidt and Simone, 1996). By studying articulation work and how to support it using digital tools, Schmidt and Simone (1996) proposed the concept of coordination mechanisms: a construction composed of a protocol instrumented by artifacts. The protocol is a resource for action consisting of a set of procedures and conventions defining the coordination work between different actors. Coordination mechanisms describe the way practitioners reduce the complexity of articulation work (Schmidt and Simone, 1996).

Coordinative artifacts are not necessarily digital; there are many examples of tangible coordination artifacts (e.g. kanban, form, checklists, etc.). Their materiality offers a better affordance and malleability but present constraints in terms of updating and visibility when they are shared. To overcome these limits, Schmidt and Simone (1996) proposed the concept of computational coordination mechanisms (CCM) that aim to digitize a part of the protocol and offer a symbolic support of the protocol through a digital artefact.

Schmidt and Simone (1996) raise two main issues regarding the design of CCM: malleability and connectivity. Indeed, practitioners need to shape and reshape the CCM in order to adapt it to the evolution of the protocol (over time or punctually) to local practices and to different contexts. As for connectivity, a CCM must be connected to other information systems, in particular business information systems and other coordinative artifacts.

In the case of large-scale coordination, Schmidt has proposed the ordering system concept. There are systems that bring order to the collection of artifacts used to support coordination, as shown in the practices of an architectural design office (Schmidt and Wagner, 2004). This study showed that, in this office, coordinative practices were organized into clusters of artifacts that make it possible to organize production and monitor changes. Overall, they reduce the complexity of coordination. For example, a drawing system is a cluster made up of a catalog of components, detailed drawings and a list of drawings, all these artifacts being linked together by references. It is the conjunction of these artifacts that allows the actors to have an overall view of the cooperation process, to know where they are and what remains to be done. This reduces the complexity of scheduling and monitoring the project.

We can identify that in the conceptualization of the digital twin there is a tension between Grives' model and Tao's model. The first model aims at proposing a single virtual space. The second model aims at breaking down Grives' model into a multitude of spaces, based on the data federating the services, the virtual replica and the physical system.

4 The overall organization of the work scheduling activity

In this section, we present the work scheduling activity that we have observed.

First, work is performed on a complex and dangerous system and in a tense economic context. Site workers are highly exposed to two major occupational hazards that can lead to fatal injuries: impact risk and electrical risks. To avoid these risks, it is necessary to cut off or reroute traffic from the work area. Then procedures define how to work in a safe manner. There are two main types of procedures: S9 procedures to remedy the risk of collision and S11 procedures to remedy electrical risks. The application of these procedures implies action from site workers and switchmen.

In addition, there are two types of work: development work and maintenance work. Development work aims to develop the capacity of a section's infrastructure (increase the number of trains per hour, safety or speed). They are major works carried out over several months or even several years. The maintenance works aim to maintain the capacities of the system in the face of its natural deterioration (due to its use) or exceptional (storms, falling trees, vandalism, etc.). Maintenance works are small or medium scale (a few days, a few weeks). Maintenance work also includes infrastructure monitoring.

To ensure the safety and the execution of maintenance and development work, three establishment types are involved:

- *Infrapoles* are in charge of the maintenance work management in their geographical sector. They are responsible for the technical safety of the train traffic (e.g. if a train derailed due to the infrastructure state; the local Infrapole is legally accountable).
- *Infralogs* are in charge of the development works execution in a geographical sector. They manage the organization of several development works. They are the interface between external railway construction companies (who carried out the works) and SNCF Réseau (who define how it should be done).
- *Railway traffic establishments* manage the train routes, regularity and safety over a geographical sector.

Over this structure, we have observed work scheduling teams that lead the work scheduling process. They are co-directed by the traffic establishment and the infrapoles and they are made up of two task forces:

- The anticipation task force, which defines an anticipation, schedule from two years before to eight weeks before the work.
- The coordination and scheduling center (called coordination center in the remaining parts) which schedules the works during the pre-operational phase from eight weeks to the day before the works begin. They adapt the anticipation schedule according to the context.

In the following, we will use the notations Y-X, W-X, D-X (with X a number), for indicating respectively X years, weeks, days, before the year, week, or day the work will be carried out.

Finally, the work scheduling activity is a long-term process. It starts two years before the work, up to the day before the work. Our study focuses on the pre-

operational phase, from W-8 to D-1. This phase is interesting because it requires a lot of articulation work among actors who must coordinate to adapt the anticipation schedule according to hazard or schedule modification (e.g. late work). The process starts at W-8; from W-8 to W-6 work requests are sent to the coordinator in the coordination center. It ends the day before the work with the editing of a daily work notice sent to the switch posts.

A key moment of the process is the coordination meeting at W-6, chaired by the coordinator. During the meeting, the requesters and the coordinator discuss the schedule, the incompatibilities, and the co-activities. They deal with unexpected events in order to "work things out" (Corbin and Strauss, 1993).

5 Method

Motivated by understanding how we could make use of the railway infrastructure's digital twin to support the coordination among stakeholders of the work scheduling activity, we have collected data through an ethnographic study in two coordination centers.

We started with two interviews with two managers of the work scheduling team in the first center. Our discussions with a strategic director of work scheduling activity enabled us to identify that the practices of this center were different from other coordination centers. We therefore decided to extend our study to a second coordination center. Our study aims to identify the coordination mechanisms involved in the work scheduling activity during the pre-operational phase and to design a technological probe (Hutchinson et al., 2003) to identify the potential use of the digital twin to support coordination in complex situations. We began by identifying when the articulation work was the most intense and focused our analysis on the coordination protocols and the artifacts in use.

We report here data collected between February and December 2020:

- Center 1: two interviews with two managers, one interview with an applicant, and observation of a 2-hour coordination meeting.
- Center 2: an interview with a coordinator, three interviews with three applicants and the observation of three coordination meetings for a total duration of 3 hours 40 minutes.

The first author conducted all the semi-structured interviews. Due to the COVID-19 pandemic, two interviews were conducted online, and all the meetings were observed online. During the interviews, we collected documents and took pictures and videos of artifacts in use. Subsequently, the interviews were transcribed. The transcripts and the materials collected were analyzed, paying attention to conventions and artifacts. We then compared our results from the two observed centers and highlighted the differences and the similarities between the practices in these two centers. The analysis was conducted by the first author, then shared and amended by the two others. Our analysis allowed us to identify three coordination mechanisms, on which we have based the design of a technology

probe to try out the role of the digital twin and to challenge our understanding of the work scheduling practice (through the use of the probe by practitioners in real settings).

6 Scheduling the Maintenance Work: A Cooperative Activity

Our fieldwork has highlighted a network of protocols and a cluster of artifacts used to coordinate work requests. We first briefly expose the observed protocols and artifacts, and focus on the detection of scheduling clashes.

6.1 The Protocol

First, to discuss the scheduling of works during the coordination meeting, coordinators build a pre-schedule. In center 1, the pre-schedule is made up of a local schedule constructed by scheduling assistants in infrapoles and infralogs. In center 2, the pre-schedule is made up of work requests sent by local managers in infrapoles and infralogs. The pre-schedule is sent to requesters (scheduling assistant in center 1, local managers in center 2) the day before the coordination meeting (at W-6).

Then, the coordination meeting chaired by the coordinator and gathering together all the requesters takes place. In center 1, the coordination meeting is about work requests at W-6, whereas, in center 2, the coordination meeting is about work requests at W-6, W-3, and W-1. In both centers, the pre-schedule is reviewed. To do so, the coordinator checks each request: the information composing it, details about the work execution, and the scheduling.

During the coordination meeting, the coordinator also resolves scheduling clashes. Indeed, if two works are scheduled at the same place and at the same moment, the requester must find an arrangement. If the works are compatible, they can be done together, but special procedures must be set up: this is called co-activity. If the works are incompatible, a trade-off has to be found or the coordinator's hierarchy will arbitrate.

After the coordination meeting, the coordinator sends a final schedule to the requesters. From W-6 to D-1, changes occurring on the final schedule due to hazard or work rescheduling. Finally, at D-1, the coordinator sends a daily work notice to switch men in switch posts. In the end, the daily work notice is less than 80% reliable in both centers.

6.2 The Artifacts

Scheduling practices use different kinds of documents.

First, S9C and S11 diagrams that respectively describe the protection procedure that has to be followed before work on the tracks or the catenaries takes place. S9C

diagrams (figure 1) indicate the elementary protection zones (ZEP) for each station and line. ZEPs are track sections where the work can take place, and S11 defines elementary catenary sections (SEL) where work can take place.

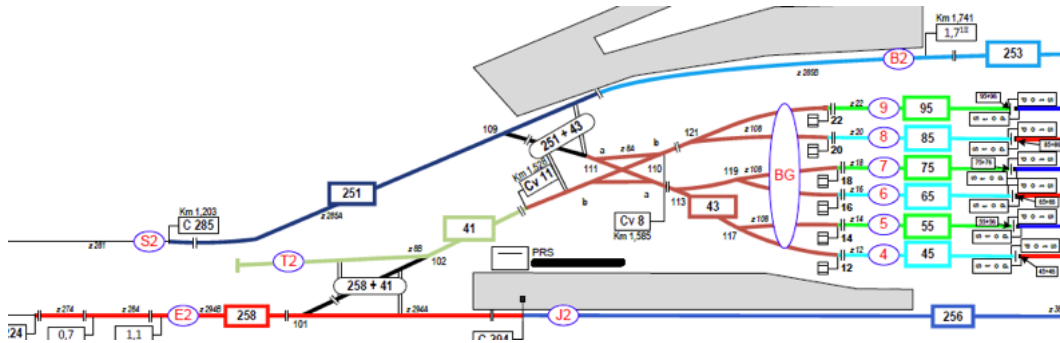


Figure 1. Extract of S9C diagram. Colored tracks show the ZEP areas, with ZEP numbers.

On top of S9C and S11 diagrams, practitioners use temporary work instructions, which are local and temporary S9C or S11 instructions used to overcome some complex situations.

Finally, the coordinator uses three other artifacts:

- Corse: An application from the information system to manage work requests.
- Spreadsheets called local schedule, pre-schedule summary or final schedule summary (depending on the scheduling state). The spreadsheets list all the work requests for an area, with details for each request.
- Coordination diagrams: S9C diagram with work areas being colored (figure 2).

Those three artifacts mobilize the S9C or the S11 information.

In both centers, the schedule summary and the coordination diagrams are used to support the articulation work. Initially, Corse was designed to support articulation work between requesters and coordinators, but we have noticed how the coordinator is using spreadsheets to workaround Corse in order to bypass the built-in workflow.

6.3 The clash detection

Clash detection is one of the main issues of the work scheduling activity. In order to prevent scheduling clashes during the coordination meeting, the coordinator must identify them before. Clashes are defined as two works scheduled at the same time and place.

Clash detection relies on the pre-schedule summary spreadsheet, the S9C and the temporary work instructions. To detect a clash, the coordinator identifies the work area of each request. To do so, they review each work request listed in the pre-schedule summary (spreadsheet). Indeed, for each work request, the day, week, starting and ending times are indicated. ZEPs, ZEPs' groups, switch posts,

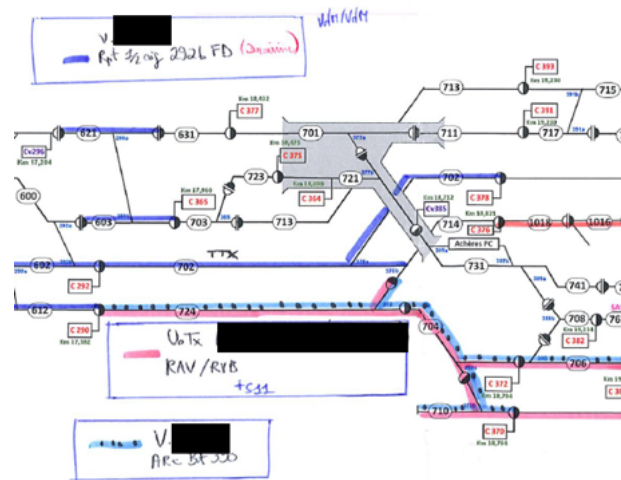


Figure 2. Extract of Coordination diagram. Colored tracks show the work areas, with ZEP numbers in the oval.

elementary catenary sections, and stations also help to identify where works are scheduled.

Therefore, in order to identify clashes, the coordinator looks, for each night, the ZEP or the ZEPs groups that are planned to be used during the night. If it is a ZEPs group, the coordinator breaks it into ZEPs, following the S9C. The coordinator then colors the work areas. If he identifies two works in the same area on the diagram, he then raises a clash. The coordination diagram identifying clashes is then sent to the requesters before the coordination meeting in which the clashes will be discussed.

7 Technology probe

The analysis of our fieldwork based on coordination mechanisms has allowed us to identify that the scheduling activity makes use of a cluster of five main artifacts (S9C, S11, Corse, schedule summary, and coordination diagram). When focusing on the clash detection practice, we identify that coordinators use the schedule summary, S9C, and coordination diagram as follows:

- The schedule summary lists all work requests, ZEP name identified work areas
- The S9C defines the ZEP localization and associates a name for each ZEP
- The coordination diagram, made by the coordinator, transcribes the schedule summary into a diagram using the ZEP definition in the S9C.

To study if we can design end-user applications supporting coordination based on the railway infrastructure's digital twin, while continuing to understand the work scheduling practices, we have designed a technology probe that is a computational version of this identified coordination mechanism. This CCM would automatize the

clash detection activity and let the human actors focus on resolving the conflicts. Regarding the symbolic aspect of the artifact, we have tried to design an artifact close to the coordination diagram that is currently used. However, track diagrams cannot be easily digitalized, so we have chosen a cartographic interface to depict work areas and clashes.

The input data are the digital twin data (that contains information that is in the S9C), the work summary and pre-schedule summary spreadsheets. Comparing data from the digital twin with the pre-schedule summary spreadsheet allows the identification of areas of work requests and therefore clashes.

The probe consists of five components:

- A script to structure ZEPs and ZEP's group data. It initiates a local cache database that stores ZEP data from requests on the digital twin into a cache database structured by switch post and station. It initiates the probe.
- A script to automatically translate work summaries into geographical data. It takes the work summary and the cache database as input and for each request: It identifies the switch post or the station, then the ZEPs or ZEP's groups involved in the request, searches for the geographical ZEP features in the database and associates request information to the ZEP shape. Finally, the script associates all the requests into a GeoJSON work summary (request.geojson) where each work request is described by a geographical feature.
- A script to automatically translate work summary into geographical data. It takes the GeoJSON work summary as input and makes a geographical intersection between work requests. If the intersection is not null, a clash object is created with the shape of the intersection and the properties of both conflict requests. All the clashes are aggregated into a clash.geojson
- A cartographical platform to visualize work requests and clashes on a map. Both files, "request.geojson" and "clash.geojson" are interpreted by the ArcGIS online platform (Esri, 2020). By using the web app-builder, we have made a web application that shows a map with requests and clashes.

The application then allows the visualization of request and clash details, to filter requests and clashes by day, and to navigate through the requests in a table. (figure 3)

To update the cartographical representation, the coordinator has to run the second and third script to produce an updated "request.geojson" and "clash.geojson" that he can drop on the cartographical platform.

This probe demonstrates two things: firstly that the data contained in the SCNF Réseau digital twin are compatible with the ones that are contained in the artifacts that support the actual practice (S9C and ZEP). Secondly, building end-users applications on the digital twin can benefit practitioners. Indeed, in our case, the coordinator puts a lot of effort into coloring the coordination diagram, based on all the requests. The probe, which integrates the digital twin's data (S9C and ZEP), can then link the schedule summary and the coordination diagram and therefore automate clash detection.

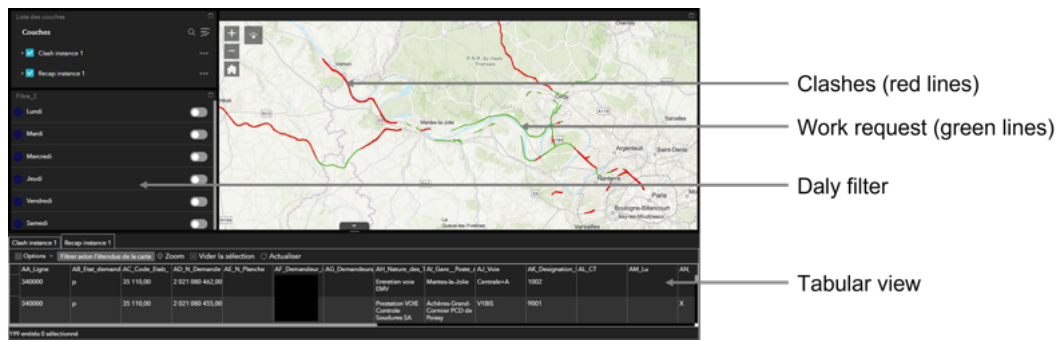


Figure 3. Cartographic probe user interface.

8 Discussion

Our work provides two contributions. The first consists of demonstrating the interest of a digital twin to rapidly deliver an application that supports coordination practices that were not easily supported before; In other words, the digital twin probe is a starting point towards an information system supporting effortless coordination (Schmidt, 2002). Indeed, thanks to the existing data model, data warehouses, and the updating process, we were able to deliver the probe with the definition of 1,508 ZEPs and 605 ZEP's groups. The needed data was requested with 35 HTTP requests. By relying on these technological capacities, and with the help of a cartographical platform, we were able to design the probe that shows requests and clashes over three months. To go further, we have identified two key factors enabling the implementation of the CCM. Firstly, as the semantics of the digital twin have been co-constructed with the practitioners, the objects, their definitions and their links are consistent with the rail network elements' technical definitions and with the terminology used by the practitioners. Secondly, as defined by Tao and Zhang (2017), the digital twin is a physical, virtual system and a set of services evolving around a common data set. In the SNCF Réseau information system architecture, this translates into data warehouses requested by several services. Thus, the data is normatively defined (by the data model) but its use is not pre-defined. Getting a global view of the infrastructure just requires requesting a unique data repository. The data warehouses then significantly simplified the development of the probe.

Our second contribution is related to the potential of the digital twin for infrastructuring. Indeed, the technology probe we have designed bridges the different coordination artifacts through the data it embeds, which profoundly transforms the information system from an ensemble of disconnected applications and artifacts to an ordering system and towards an information infrastructure. This finding supports the conclusions of Susan Star (1999), Bossen and Markussen (2010): digitization makes it possible to automate the links between the artifacts used by the practitioners, which reinforces the coupling between the artifacts but

also makes the links between them invisible, which tends towards the construction of an information infrastructure. We therefore claim that the digital twin is a step towards the infrastructuring of the ordering system.

9 Conclusion

The digital twin is one of the technologies at the heart of Industry 4.0. Even if one of the intended uses of the digital twin is to support the cooperation between different businesses by providing a shared digital representation, few studies have investigated this opportunity.

Thus, in this exploratory study, we sought to examine the role the digital twin can play to support articulation work. We conducted a multi-sited ethnography in two work scheduling centers, which allowed us to highlight the entanglement of practices and artifacts that coordinators and requesters mobilize. Based on this study, we have designed a technological probe based on the digital twin. The results of this study firstly show that the work scheduling practice requires complex articulation work. Secondly, practitioners articulate around a constellation of artifacts, and the links between these artifacts are performed manually. Thirdly, we quickly achieved the design of a technological probe that supports a part of the protocol. Fourthly, this probe highlights that the digital twin tends to automate the links between artifacts. Finally, we argue that the digital twin participates in the creation of an information infrastructure. In this infrastructure, we have identified the importance of referential data: data shared between different professions, normatively considered as true.

Acknowledgements

We thank the coordinators and the managers of the coordination centers, unit managers, local managers, assistants, and all the other actors of SNCF Réseau for their participation in this study. We also thank Eddie Soulier for his management of the UTT/SNCF Réseau research program. Finally, we would particularly like to thank the actors of the digital twin at SNCF Réseau, in particular Bruno Landes, Sofia Costa D'Aguiar and Jean-Christophe Michelin who made this research possible. This research is funded by SNCF Réseau in the framework of the BIM and digital continuity program.

References

- Bjørn, P., M. Wulff, M. S. Petrus, and N. H. Møller (2021): 'Immersive Cooperative Work Environments (CWE): Designing Human-Building Interaction in Virtual Reality'. *Computer Supported Cooperative Work (CSCW)*.
- Bossen, C. and R. Markussen (2010): 'Infrastructuring and Ordering Devices in Health Care: Medication Plans and Practices on a Hospital Ward'. *Computer Supported Cooperative Work (CSCW)*, vol. 19, no. 6, pp. 615–637.

- Corbin, J. M. and A. L. Strauss (1993): 'The Articulation of Work Through Interaction'. *The Sociological Quarterly*, vol. 34, no. 1, pp. 71–83.
- Costa D'Aguiard, S., B. Landes, G. Ducelier, and C. Stalder (2019): 'Digital twin: concept and first developments into the management of a cyber-physical railways system'.
- Esri (2020): 'ArcGIS Web AppBuilder | Documentation'.
- Gartner (2017): 'Gartner Top 10 Strategic Technology Trends For 2018'.
- Gautier, P. E., S. Costa D'Aguiard, B. Landes, and C. Stalder (2019): 'Vers un Jumeau Numérique du Système Ferroviaire'. Livre blanc, SNCF Réseau, Paris.
- Grieves (2015): 'Digital Twin Manufacturing Excellence through Virtual Factory Replication'. Technical report.
- Guinet, A. and H. De Nicolay (2019): 'Rapport Financier SNCF RESEAU 2019'. RAPPORT FINANCIER DU GROUPE SNCF RÉSEAU, SNCF Réseau, La Plaine Saint-Denis.
- Hutchinson, H., W. Mackay, B. Westerlund, B. B. Bederson, A. Druin, C. Plaisant, M. Beaudouin-Lafon, S. Conversy, H. Evans, H. Hansen, N. Roussel, and B. Eiderbäck (2003): 'Technology Probes: Inspiring Design for and with Families'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 17–24, Association for Computing Machinery. event-place: Ft. Lauderdale, Florida, USA.
- International Railway Solution (2016): 'IRS 30100: RailTopoModel - Railway infrastructure topological model'. IRS 30100: RailTopoModel - Railway infrastructure topological model IRS 30100.
- Lamb, K. (2019): 'Principle-based digital twins: a scoping review'. Technical report, Apollo - University of Cambridge Repository.
- Ministère de l'économie et des finances (2019): 'Décret no 2019-1587 du 31 décembre 2019 approuvant les statuts de la société SNCF Réseau et portant diverses dispositions relatives à la société SNCF Réseau'.
- Møller, N. L. H. and J. P. Bansler (2017): 'Building Information Modeling: The Dream of Perfect Information', pp. 18.
- Nochta, T., L. Wan, J. M. Schooling, and A. K. Parlikad (2021): 'A Socio-Technical Perspective on Urban Analytics: The Case of City-Scale Digital Twins'. *Journal of Urban Technology*, vol. 28, no. 1-2, pp. 263–287.
- Pires, F., A. Cachada, J. Barbosa, A. Moreira, and P. Leitao (2019): 'Digital twin in industry 4.0: Technologies, applications and challenges'. Vol. 2019-July. pp. 721–726.
- Schmidt, K. (2002): 'The Problem with 'Awareness': Introductory Remarks on 'Awareness in CSCW''. *Computer Supported Cooperative Work (CSCW)*, vol. 11, no. 3, pp. 285–298.
- Schmidt, K. and C. Simone (1996): 'Coordination mechanisms: Towards a conceptual foundation of CSCW systems design [1996]'. *Computer Supported Cooperative Work (CSCW): The Journal of Collaborative Computing*, vol. 5, no. 2-3, pp. 155–200.
- Schmidt, K. and I. Wagner (2004): 'Ordering Systems: Coordinative Practices and Artifacts in Architectural Design and Planning'. *Computer Supported Cooperative Work (CSCW)*, vol. 13, no. 5, pp. 349–408.

- Simone, C., K. Schmidt, P. Carstensen, and M. Divitini (1996): 'Towards a technology of coordination'. Publisher: Citeseer.
- Star, S. L. and A. Strauss (1999): 'Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work'. *Computer Supported Cooperative Work (CSCW)*, vol. 8, no. 1-2, pp. 9–30.
- Susan Star (1999): 'The ethnography of infrastructure.'. *American Behavioral Scientist*, vol. 43, no. 3, pp. 377–391.
- Tao, F., J. Cheng, Q. Qi, M. Zhang, H. Zhang, and F. Sui (2018a): 'Digital twin-driven product design, manufacturing and service with big data'. *The International Journal of Advanced Manufacturing Technology*, vol. 94, no. 9-12, pp. 3563–3576.
- Tao, F., Q. Qi, L. Wang, and A. Nee (2019a): 'Digital Twins and Cyber–Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison'. *Engineering*, vol. 5, no. 4, pp. 653–661.
- Tao, F., H. Zhang, A. Liu, and A. Y. C. Nee (2019b): 'Digital Twin in Industry: State-of-the-Art'. *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2405–2415.
- Tao, F. and M. Zhang (2017): 'Digital Twin Shop-Floor: A New Shop-Floor Paradigm Towards Smart Manufacturing'. *IEEE Access*, vol. 5, pp. 20418–20427.
- Tao, F., M. Zhang, Y. Liu, and A. Nee (2018b): 'Digital twin driven prognostics and health management for complex equipment'. *CIRP Annals*, vol. 67, no. 1, pp. 169–172.

Mareike Glöss (2022): Connectedness in mobile families. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_ep09

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_ep09

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

Connectedness in mobile families

Mareike Glöss

Malmö University, Sweden

KTH, Royal College of Technology, Sweden

mareike.gloss@mau.se

Abstract. Family life is no longer confined to geographically shared spaces. More often, families are separated. Technology offers countless means of keeping families connected, which has been subject of extensive research. Yet, connection between families goes beyond interpersonal communication. Being separated from extended family means to be separated from familiar rituals, habits, and values. In this paper we present an ethnographic study of mobile families to understand how families are dealing with this kind of separation in their everyday life. We analyze situated practices and discuss how these families create a sense of connectedness to their country of origin. Our observations show that design for connectedness should address practices and materialities that are part of the family home. Furthermore, we argue that there should be more consideration for what the family connects to: Instead of connecting between people, connectedness can also be seen as staying in touch with familiar routines, customs, and environments.

1. Introduction

Social interactions within families rely increasingly on digital technologies. Urbanization, globalization, ICT, and modern economics lead families to relocate to a new or temporary home, often in distant and foreign environments. And the past two years of the global COVID-19 pandemic have shown how resourceful

families can be in maintaining the social fabric of modern family life – even when forced to social distancing. Digital technologies open a wide range of opportunities for families to stay connected, however, the applications and devices used have remained remarkably static. Connectedness between families is still mostly limited to different forms of communication, mainly through sound and video transmission. But families that are separated from their extended family and close friends are also experiencing a separation from familiar values, rituals, and habits. Often, families are challenged to create a sense of connectedness while at the same time making a home.

There is a wide body of research on ways to support emotionally rich distant communication between family members. Yet especially design-led inquiries are mostly focused on inter-personal instances of connection. Less attention is paid to the role of connectedness as a continuous experience, present in the many everyday practices in family life. In this paper we argue that it is necessary to revisit and expand existing research on interpersonal connectedness. Hence, we scrutinize connectedness not just in terms of concrete instances of communication but as embedded in the rich practices of everyday life. Thus, we are seeking to explore how families are creating this particular experience in their everyday, for instance, through the use of objects, through arranging spaces or by maintaining ritualized routines. We argue that both, practices as well as the artefacts that are part of these practices, play a crucial role in the emergence of connectedness. And to understand these experiences we must dissect the practices and the material culture behind them.

To make the practices of connectedness more visible we decided to study an extreme group: We focused on families that are often months or years separated from their originate home, through country or even continent borders. We report on an ethnographic study that was conducted before the outbreak of the Covid-19 pandemic. We studied the everyday life of ten families that live outside of their home countries – we call them here ‘mobile families’. For the researcher in the field, the contrast between local and foreign artefacts and practices facilitated identifying and tracing practices of connectedness throughout the family’s everyday life. Yet, we argue that, while the families presented here must overcome quite high obstacles, the challenges they are dealing with, also give insights into practices of all families that have to deal with any kind of spatial and social distances.

In this paper we want to make three contributions: First, we want to show the variety of ways in which families integrate the sense or experience of connectedness into everyday practices. Second, we show some of the elements and mechanisms around these instances of connectedness. Lastly, we show the implications of this experiential view on connectedness for the design of new technologies that can enhance socially distanced family life.

2. Background

Connectedness, in particular in regard to families, is a well-researched topic in HCI. While earlier work has been dealing mostly with connectedness at the workplace, the recent decade has seen increased interest into connectedness at home and among families. Here, the term has found reflection mainly in its definition as “interpersonal connectedness” and the relationship between social contacts (Van Bel et al. 2008). Research in this field has thus mainly focused on the communication between people, for instance in long-distance relationships (Alsheikh et al. 2011) or in a migration context (Wyche and Chetty 2013).

We want to introduce the most relevant strands of research on connectedness, beginning with some of the conceptual discussions around presence followed by a review over research and design on families and connectedness.

2.1 Presence and awareness

Presence and awareness have become prominent terms in order to describe the experiential dimension of interpersonal connectedness. The work done in this area illustrates the challenges of grasping connectedness on a conceptual level. The emerging expanded understanding of connectedness has influenced designs that aim to bridge the distance for instance between couples in long-distance relationships (Mueller et al. 2005; Lottridge et al. 2009; Hassenzahl et al. 2012) or between older people and their children and grand-children (Mynatt et al. 2001; Durrant et al. 2009; Wallbaum et al. 2018). Several design studies have focused on the experiential qualities of connectedness and go beyond verbal communication but instead create connection through bio-signals (Min and Nam 2014), connected lighting (Morris et al. 2017) or the transmission of pressure (Mueller et al. 2005) Others focus on shared activities, such as meals (Grevet et al. 2012) or watching TV (Harboe et al. 2008). There exist many interesting design explorations in this area, for instance Tibeau et. al.’s ‘Family Song Box’ (Tibau et al. 2019).

Early on in the design of different communication technologies, such as video displays, the concept of presence served as theoretical foundation e.g. (Bly et al. 1993). Originating from the field of social psychology (Short et al. 1976) it emphasizes the importance of non-verbal communication, i.e. communication not expressed merely verbally. A wide range of design explorations have since aimed at created this sense of presence with the help of technology (Bly et al. 1993; Kuwabara et al. 2002; IJsselsteijn et al. 2003).

But with technology entering the home, HCI research turned towards a more complex understanding of connectedness. Rettie (2003) argues that connectedness does not necessarily require direct social contact but that the focus should instead lie on the overall experience. While presence might describe an experience during

a specific instance of communication, several studies and design explorations focus on a more implicit form of connectedness that is not temporally limited but continuously present. For instance, Bales et al. (2011) designed a system that automatically shares the location between couples with the help of rather unobtrusive cues. Others have integrated subtle forms of information exchange into everyday objects (Dey and Guzman 2006), connect households by transmitting traces of touch or objects via a table (Dey and Guzman 2006), or allow for digital storytelling between parents and their children (Wong-Villacres and Bardzell 2011; Cheong and Mitchell 2015).

As a term, awareness has emerged as an understanding of connectedness as being aware of each other or “to keep in touch with” (p. 187) others. (Liechti and Ichikawa 1999). This line of research focuses primarily on inter-personal connection. Lynggaard et al. (2010) as well as Ylirisku et al (2016) take a slightly different path and explore connectedness in relation to places. Both studies are conducted by placing particular artefacts into family homes that connect to a different personal place like the family’s summerhouse, through image or video material. Thus, they create what Ylirisku et al refer to as “place-presence”.

2.2 Connectedness and Family life

Families and the family home have emerged as an important subject of HCI research and as target for a variety of design explorations. Here, materiality and practice as analytical stances had a big influence; for instance, in form of research dealing with the importance of material artifacts in supporting a sense of domestic connectedness and family relations (Taylor and Swan 2005; Wakkary and Maestri 2007). Bales & Lindley (2013) look at how particular artifacts in the lives of undergraduate students convey the character of the parental home. And Odom et al. (2010) describe the material infrastructure of possessions and arrangements that characterize the bedroom of children living in divorced homes. A wide range of studies has explored how families can be connected to each other. Some of these studies focus on connectedness between families living apart from each other (Ames et al. 2010; Brown and Grinter 2012) while others consider the connectedness of the nuclear family (Brush et al. 2008; Durrant et al. 2009). These studies are mainly focused on moments of family interaction and communication. Other studies address the family home, awareness and the sensorial connection to place (Dey and Guzman 2006; Brown, et al. 2007; Lynggaard et al. 2010) and have also found reflection in the form of ambient systems (Agamanolis 2005). This line of research shows the importance of digital technologies in interpersonal connectedness and especially because they show how technology enhances family life and contributes to well-being. We want to expand on this research by focusing on connectedness as it is cultivated and maintained in practices.

3. Methodology

The data presented here draws on the results of an ethnographic study carried out in the Netherlands with ten families over the course of two months. Because of our interest in families' individual experience and perception of their everyday life, we chose a phenomenologically informed approach, blending elements of semi-structured interviews with observations. Our main interest were mundane practices and routines that were in some way related to the subjects' home countries.

Within anthropology, there are a variety of methodological approaches that seek to get as close to the individual experience of a participant. We designed this study loosely based on Kusenbach's (Kusenbach 2003) go-along approach. This approach has the researcher walk with the participant and share their experiences (in Kusenbach's case outside) while at the same time reflecting on them together with the participant. Thus, this method addresses a common challenge in studying everyday practices and individual experiences: Any kind of routine or interaction that is part of everyday life is usually taken for granted by participants. Walking along with the families in their homes made it possible to make the apparently mundane visible.

Another important pillar of our research methodology is the concept of practice as an analytical frame for understanding activities and interactions. In the last decade, HCI research has demonstrated a growing interest in social theories of practice (Kuutti and Bannon 2014) and a renewed interest in technology's role in everyday actions. This emerges out of the desire to untangle the increasingly complex and diverse practices that are emerging around digital technologies (Goodman and Rosner 2011).

In this part of the paper, we describe the participants, the overall setup and procedures during our observational interviews as well as our analysis procedures.

3.1 Study setup

Families were visited at home, usually on the weekend at a time when all family members were present. The visits lasted approximately 3 to 4 hours each, during which the ethnographer would also participate actively in a family activity, such as dining or going for a walk. In the beginning of the visit, the researcher explained details about the study and procedures around any collected data and answered participants' questions. This was followed by a detailed tour of the house in which the researcher would repeatedly ask for certain artefacts or practices (e.g. "When was the last time you have used this desk?" "Can you tell me more about this picture?") Keeping the visit informal and interactive was a priority and the researcher would actively engage in different activities, such as

(playing with the children or help with meal preparation. The visits were recorded. Additionally, we took approximately 80-100 photos per family and shot short videos of some families performing certain practices. This material was later used as cues and mementoes during the analysis process.

Table 1: Overview over participating families

Alias	Family Members and countries of origin
Family A	Mother, father and one son (6 months) – from Indonesia
Family B	Mother (Uganda), father (Kenya) and one daughter (2 years)
Family C	Mother, father, one daughter (8), one son (5) - Columbian
Family D	Mother (USA), father (Germany), one son (8) one daughter (4)
Family E	Mother (Brazil), father (Hungary), one son (3)
Family F	Mother (Italy), father (Kenya), one daughter (3)
Family G	Mother (Sweden), father (France), daughter (3), son (1)
Family H	Mother (UK), father (Dutch), son and daughter (both 2)
Family I	Mother (USA), father (German, but grew up in USA), one daughter (3)
Family J	Mother and father, two sons (3 and 1) - Dutch

3.2 Participants

Ten families were enlisted for this study. We recruited nine of these families mainly through Social Media groups related to expat life. Thus, while the participants' personal backgrounds as well as duration and reasons of stay abroad differ, they all do identify themselves as expats or mobile families.

The tenth family was a Dutch family without any migration background, recruited through our personal network. This family served as a control case – in the widest sense of the word. Contrasting our observations with the non-migrant family to our observations in other families helped us in identifying those practices that are primarily related to family life rather than connectedness. All families received compensation for their participation.

The background of the families as well as their socio-economic situation is very diverse. In some cases, both partners were from the same country and had

moved together to the Netherlands (Families A, C and I). Other partners had met in the Netherlands while studying or working and decided to build their family there (Families B and E). And some families had already been living in other countries as expats and had moved from there together to the Netherlands (Families D, F, G and H). The families also differed in their long-term plans as it related to the Netherlands; some families had decided to settle down there, while others were in a more transitional state and were planning to move back to their home countries or to another country soon.

3.3 Data Analysis

In preparation for the analysis, all interview recordings were transcribed and anonymised. Furthermore, the field researcher gave a subjective account of her experiences and described the family's everyday life based on her observations. These thick descriptions of family life were written on the same day that the observation took place. All collected data – audio transcripts, photo- and video material and the family descriptions – were then coded using atlas.ti. The main part of the analysis took part during two workshops. During these workshops, the data was examined by a team of researchers working on this study and through the lens of our interpretive resources. In these workshops we started by identifying as many practices and routines as possible, before clustering them around different themes and scrutinizing their function in relation to connectedness. Throughout several iterations we identified those practices that were directly linked to connectedness, in particular to the participants country of origin. In the following, we will present these practices and their function within the family home.

4. Observations

In our analysis, we found distinct practices emerge around artifacts, demonstrating the variety of different strategies and routines supporting the families' needs for connectedness. We introduce the most important practices that are related specifically to connectedness and exemplify them along moments or vignettes from the field.

4.1 Traveling, communicating and re-creating – bringing the home country to life

Many practices in the family's life were obvious in their connecting function: The long travels back home that come with the arduous coordination around school holidays and work; the calls back home to keep in touch with grand-parents and friends, organized by mobile families through a multitude of channels and

material arrangements. Technology played a large role in this, and it became apparent that it had already enhanced both communication as well as the possibility to receive information, movies or photos from the home country.

Maybe not surprisingly, those practices that stood out in terms of connectedness were those of nostalgia or remembrance. In many households, certain artefacts were presented to the researcher as valued treasures or just plainly as things one would “never throw away”, they were clearly an important pillar of many participants’ connectedness to their home countries:

The mother of family E shows me a special christening towel that she keeps in the drawer with the rest of the regular-use towels, even though she would never use it. The towel played an important part of her son’s christening ceremony in her home country and has been embroidered by her two aunts. The towel is stored in an everyday place thus is a constant reminder for her of this very important religious ritual, while at the same time also connects her to her family members back home and connecting her son to them and his heritage. (As shown in Fig. 1)

Those artifacts did not necessarily exhibit folkloristic elements typical to their home country. Often these were just objects or arrangements they used to have back home:

The mother of family F shows me how she has arranged jewelry cases and cosmetic equipment on the window board of her small bedroom. She has them arranged in the exact same fashion as she has in every bedroom she’s occupied since she was in her childhood bedroom in her home country. She tells me that it gives her the sense of her own place, even when space in the apartment is limited.

These two observations illustrate that many practices of connectedness are not necessarily related to communication but rather to a certain use of artefacts or spaces. The arrangement of things, the way in which certain practices are performed around certain family heirlooms or keepsakes represent other ways of staying in touch with the home country.



Figure 1 and 2: Christening towel and mantelpiece arrangement.

4.2 Storing, arranging and decorating – Making a home

One of the biggest challenges for mobile families seemed to be to make a home, i.e. to create the place they are living in in a way that makes them feel as comfortable as they would be in their home country.

The mother of family C points towards an empty wall in her living room. She tells me that she would like to decorate this wall but that she would never decorate with things acquired in the Netherlands because she does not like the Dutch taste. “I already know I will not like it, I won’t even look for anything here.” Instead, she would go all the way to Colombia in order to acquire decorative objects there.

She proceeds to show the researcher some previously acquired decorations. These items are not directly linked to Colombian heritage. These decorative artifacts are in no way linked to a particular country or region. Yet, for her it is important to express the style that she associated with her home country. Traces of this kind of home-making were among the most obvious ones during our observations. Practices of decorating or storing were often reflecting certain values or ideals and thus were expressing a certain identity. The following example shows this even clearer:

The mother of family D shows me greeting cards that are arranged all over the living room, but in particular over the fireplace. The living room is a very representative place that is seldomly used and the greeting cards have a decorative character. Families sent them from the couple’s respective home countries on special occasions, such as Valentine’s Day or birthdays. She explains to me that arranging these cards is very important to her and that she wishes that her children also value cards like these in the future instead of only focusing on digital devices. [As shown in Fig 2]

In this example the mother expresses a high appreciation not only for her family but also for postcards that represent a certain cultural meaning to her. By arranging the cards and by attempting to pass the cultural significance of them on to her children, she seeks to maintain certain values that are an important part of her connectedness practices.

Digital technologies play a lesser role in these home-making practices. While mobile phone, tablet or laptop were natural parts of the home environment and played a large role in communication, games or entertainment, they were seen as mere tools to procure the necessary materials, such as photos, movies or cooking recipes.

The here described practices stand out in that they at first seem to be singular occurrences, like decorating a mantel piece or storing a precious keepsake. Yet, during our fieldwork it became clear that most of these arrangements also involve constant maintenance. For instance, the arrangement of cards on the mantel piece is constantly changing, due to the addition of new cards. Similarly, decoration in a particular style needs to be adapted to the family’s changing circumstances, interests or tastes. And the beloved christening towel, while safely tucked away in the child’s drawer, repeatedly makes an appearance when clothes need re-arranging.

4.3 Cooking, gathering, playing – everyday family life

A third group of practices we observed were everyday routines that are at the core of each family's daily life. When looking closer at different practices it became clear that in almost every practice, families were adapting previously existing habits and routines to new life circumstances to create a sense of connectedness. While many of these practices appeared to be rather mundane at first, their importance in the families' life must not be under-estimated. For instance, all families that we interviewed had different ways of procuring country-specific toys for their children. In particular, children's' books and TV shows in the parents' native language constituted an important element of raising children in a foreign country. Books are very often procured through relatives as gifts:

The mother of family F points out some books in her daughter's library that were actually her own childhood books that her own mother had sent from Italy. This mother explains that she likes that her daughter also reads the same books she read as a child.

Examples like this were found in every single family. Raising the children according to one's own values is one of the most important elements and challenges of mobile family life. Material and non-material artifacts, like TV-shows, played a crucial role in this. Another example shows the role of digital devices:

The mother and father of family D point out the stereo system that they have installed throughout the whole house that is connected to Internet radio. Both are very enthusiastic about it because it allows them to listen to their favorite radio stations. He would listen to a German station, often using headphones, while she, coming from the US, listens to American stations during everyday activities. "Sometimes when I am in the house, I feel that I am in America. The radio is American, I am cooking American food... Only when I step out of the door, I realize I am in the Netherlands." [Vignette 4]

In this example both, mother, and father, surround themselves with digital technologies this way slightly aligning their home to what they know from their home countries and their preferred taste. The crucial role of digital technologies in the flow of our families' practices becomes also apparent in this example:

The mother of family G has a very special strategy to keep her children occupied. Sometimes, when she has some housework to do, she will call her parents in Sweden over video call. The two children are used to meeting their grandparents this way from very young age so the grandparents can function as 'babysitters' and keep the children occupied. [Vignette 7]

In this example the grandparents – or rather their visual and audio representation on the computer screen – become virtual actors. The mother needs them to watch the children – just like close-by grandparents might do – but because of their remoteness she has to connect to them via the computer. The values at play here are those of parenting strategies and keeping the family members in touch with one another.

5. Discussion

Our observations show that families go through some length in creating a sense of connectedness. And it also becomes apparent that connectedness is not limited to particular moments of connection. Connectedness presents itself as a rather ephemeral sense or feeling that becomes visible through everyday practices. Before we can think about designing for connectedness, we want to discuss the core elements and mechanisms behind the different practices we observed.

5.1 Maintaining the family home

Surrounding oneself with familiar artefacts, teaching children certain values, and celebrating certain family routines - all these practices aim towards constituting and maintaining the family home. Every family strives to maintain a certain set of values that they also seek to pass onto their children and that are an important part of their everyday life. These *practices of maintaining* can happen in tangible forms such as preserving the particular practices of food preparation or the kitchen table as a central element of gathering the family. But in our study, we also often encountered *practices of maintaining* that are more conceptual, such as those related to child rearing and socialization. Here the need to convey certain values from the home country – a certain way of doing things – plays a central part. Values become weaved in with the routines and habits developed at the new place of living. Raising a family is a highly personal and cultural process, which for a migrant in a foreign country can lead to unique challenges in maintaining practices. And earlier studies have pointed out the important role of digital technologies in this context. (Liaqat et al. 2021) Moreover, practices of maintaining and home making are not specific to expat-families (Kirk and Sellen 2010; E. S. Bales and Lindley 2013; Light and Petrelli 2014). And looking into the home as a place in which identity and the self are expressed is not a new concept in HCI research (For instance (Swan et al. 2008; Odom et al. 2010; Ylirisku et al. 2013). Most practices of maintaining, such as the exhibition of family photos or the specific arrangement of decoration and furniture were for instance also observed in the Dutch family maintaining their family and home in their country of origin. But what distinguishes migrant families from those families that are living in their country of origin is that the materials they may require for their unique practices of maintaining may not be readily accessible. Instead, some of the families were constantly *procuring* artifacts.

5.2 Continuous procuring

On the surface the need for practices of procuring seems to be a mere practicality. For instance, because the children grow up bilingual there is a need for books in

their other language. However, examining practices of procuring further shows the deep connection to maintaining family values: The mother of family F got her old childhood books sent by her mother because she likes the idea that her daughter reads the same stories as she did as a child. And this becomes even more apparent when the father of family E expresses his concern that his son might speak the language of his home country fluently but won't understand the, as he calls it "cultural language" such as popular TV-shows or music.

Procuring is essential in order to make a home comfortable, i.e. it allows for a maintaining of values and dispositions. Practices of maintaining and procuring work together in connecting current everyday practices with values, habits, or routines from the home country, thus creating a sense of connectedness. Sometimes the procuring presents itself as a single event, the outcomes of which are then maintained throughout everyday interactions. For instance, a once bought children book will be used regularly for reading to the children.

In many other cases procuring and maintaining are enabled digitally. In the case of Internet radio or movie streaming they are happening simultaneously; the listening or watching is happening instantly. Also procuring is sometimes happening in unexpected ways. When grandparents get to babysit their grandkids via video chat, this can be seen as procuring their presence with the help of network and streaming technology. Here it becomes very apparent which role digital technologies already play in mobile families as well as many other families that creating connections to other places and people.

Above anything, the interplay between maintaining and procuring practices highlights the dynamic and continuous character of connectedness. Connectedness is not just created in brief moments of interaction, but has to be understood as being an on-going experience present in an abundance of everyday mundane practices. People are most of the time their own designers, setting up and creating moments of connectedness throughout their daily routines. What makes this particular design space so difficult, is the existing complex web of routines, habits, values and social interactions that exist in every family home. Any design deployed, will unavoidably be affected by existing practices.

6. Design for Connectedness

There is always a risk that ethnographic insights into complex experiential phenomena cannot lead to more generalisable insights that could inform the design for many. But while every family has their own individual take on how to create connectedness in everyday life, we want to highlight some of the key aspects that they all have in common and that can serve as a point of departure for a design exploration.

6.1 Design for practices of maintaining and procuring

We have already emphasized in length that connectedness should be understood as part of everyday practices instead of a practice in itself. Hence, by identifying important practices and their joints, design of digital technologies can overcome the limitation of isolated experiences and integrate better into everyday life. Here, we suggest for any design to identify and target practices of both maintaining and procuring. Digital technologies can facilitate and expand the interrelationship between these practices, for instance by procuring digital, immaterial artefacts. This is already done for example in the form of streaming music, radio, or video. Yet, we think that design should experiment with creating digital components of many other mundane practices – expected and unexpected ones. Again, we want to emphasize that we do not understand this as pure interpersonal communication. There is already an abundance of design artefacts that allow separated families to exchange mementos such as photos or hand-written notes. Practices of connectedness can also happen on a very individual level, in order to create a sense of familiarity. Hence, we suggest exploring designs targeted at both – collaborative interaction in the family as well as individual activities.

6.2 Connected to what?

The implicit nature of connectedness has been highlighted by many before: Connectedness is not just created through short distinct moments of communication but also through awareness of each other, for example by sharing everyday activities (Romero et al. 2007; E. Bales et al. 2011) or knowing about each other's whereabouts (Brown, et al. 2007). Thus, more recent approaches to connecting people and families have been expanding our understanding of how people connect. Yet, our study shows that we must add another question to our inquiries, that is what people are connecting to. In our study, being connected was not just a form of social encounter. Being connected meant to be connected to a different life, different values, routines, and material culture. Our participants always strived for moments of immersion with certain experiences from their home countries: Listening to the radio, arranging furniture and artefacts, cooking and eating – again this immersion happens through mundane everyday practices. We suggest that the design for practices of maintaining and procuring should allow for this kind of immersion into familiar cultural and social spaces.

6.3 The importance of things and places

Within the design space of the family home, materiality, i.e., the materials and artefacts families surround themselves with, play a prominent role. For instance, Ylirisku et al. (2013) evaluate tactile experiences of artifacts and how they fit into the material landscape of the family home. Studies have also shown the

importance of artefacts in the home for identity and the self (Odom et al. 2009) and how families make their home unique through practices of storing and managing (Swan et al. 2008).

Yet, when it comes to designing for connectedness, the role of existing artefacts as means of connecting is rather neglected. While many designs mimic everyday objects such as picture frames (Mynatt et al. 2001; Romero et al. 2007) or a clock (Brown, et al. 2007), they seldomly relate to existing places and artefacts in the home. Our observations showed that personal artefacts and place within the family home are a vital part of any practice of connectedness. Any design that aims at practices of connectedness should aim to consider and connect to existing artefacts and places that are part of family life.

6.4 Tying together practices, experiences and materiality

Apart from the above-described suggestions, the core of our contribution lies in the particular theoretical lens of our inquiry: One of our aims was to add to the growing design research on interpersonal connectedness, by presenting an ethnographic approach that focuses less on technological opportunities but instead engages with existing practices on a deeper level. Many of the practices that we observed, might at a first glance be regarded as irrelevant or too mundane. But when we really scrutinized these everyday activities, we were able to uncover an experiential dimension to them. Any design needs to be built on an understanding connectedness as an experience tied to practices and materialities already existing in every family.

7. Conclusion

An increasing number of families are living spatially separated from their familiar social and cultural circles. And with the restrictions on travel and human gatherings that became part of the global fight against the Covid19 epidemic, the need for connectedness becomes even more apparent. To scrutinise practices of connectedness we studied families living outside of their country of origin. We wanted to understand how these families create a sense of connectedness with their respective home countries through their everyday interactions. Our observations show that connectedness in these mobile families is not limited to temporally limited interpersonal interactions. Instead, connectedness is also a kind of feeling or sense that needs to be constantly maintained. It describes a connection not to a specific person or social group but rather to the ephemeral sense of familiarity associated to a certain cultural environment - usually the one that one has been brought up in. Our study has shown that the sense of connectedness cannot be reduced to interpersonal connections. In designing

technological tools for connecting families, we need to expand our understanding what people are connecting to and how everyday practices are involved.

References

- Agamanolis, Stefan. 2005. New technologies for human connectedness. *AMBIENT INTELLIGENCE*: 5.
- Alsheikh, Tamara, Jennifer A. Rode, and Siân E. Lindley. 2011. (Whose) value-sensitive design: a study of long- distance relationships in an Arabic cultural context. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work - CSCW '11*, 75. Hangzhou, China: ACM Press. <https://doi.org/10.1145/1958824.1958836>.
- Ames, Morgan G, Janet Go, Joseph “Jofish” Kaye, and Spasojevic, Mirjana. 2010. Making love in the network closet: the benefits and work of family videochat. In , 10. Savannah, Georgia, US: ACM.
- Bales, Elizabeth, Kevin A. Li, and William Griwsold. 2011. CoupleVIBE: mobile implicit communication to improve awareness for (long-distance) couples. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work - CSCW '11*, 65. Hangzhou, China: ACM Press. <https://doi.org/10.1145/1958824.1958835>.
- Bales, Elizabeth S., and Siân Lindley. 2013. Supporting a sense of connectedness: meaningful things in the lives of new university students. In *Proceedings of the 2013 conference on Computer supported cooperative work - CSCW '13*, 1137. San Antonio, Texas, USA: ACM Press. <https://doi.org/10.1145/2441776.2441905>.
- Bly, Sara A., Steve R. Harrison, and Susan Irwin. 1993. Media spaces: bringing people together in a video, audio, and computing environment. *Communications of the ACM* 36: 28–46. <https://doi.org/10.1145/151233.151235>.
- Brown, Barry, Alex S. Taylor, Sharam Izadi, Abigail Sellen, Joseph “Jofish” Kaye, and Rachel Eardley,. 2007. Locating family values: a field trial of the whereabouts clock. In *Proceedings of the 9th international conference on Ubiquitous computing*, 354-371. Innsbruck, Austria: Springer-Verlag.
- Brown, Deana, and Rebecca E. Grinter. 2012. Takes a transnational network to raise a child: the case of migrant parents and left-behind Jamaican teens. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12*, 123. Pittsburgh, Pennsylvania: ACM Press. <https://doi.org/10.1145/2370216.2370237>.
- Brush, A.J. Bernheim, Kori M. Inkpen, and Kimberly Tee. 2008. SPARCS: exploring sharing suggestions to enhance family connectedness. In *Proceedings of the ACM 2008 conference on Computer supported cooperative work - CSCW '08*, 629. San Diego, CA, USA: ACM Press. <https://doi.org/10.1145/1460563.1460661>.
- Cheong, Kakit, and Alex Mitchell. 2015. Kwento: Using a participatory approach to design a family storytelling application for domestic helpers. In *IFIP Conference on Human-Computer Interaction*, 493–500. Springer.
- Dey, Anind K, and Edward S De Guzman. 2006. From Awareness to Connectedness: The Design and Deployment of Presence Displays: 10.

- Durrant, Abigail, Alex S. Taylor, David Frohlich, Abigail Sellen, and David Uzzell. 2009. Photo Displays and Intergenerational Relationships in the Family Home. In Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology, 10–19. BCS-HCI '09. Swindon, GBR: BCS Learning & Development Ltd.
- Goodman, Elizabeth, and Daniela Rosner. 2011. From garments to gardens: negotiating material relationships online and “by hand.” In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2257–2266. Vancouver BC Canada: ACM. <https://doi.org/10.1145/1978942.1979273>.
- Grevet, Catherine, Anthony Tang, and Elizabeth Mynatt. 2012. Eating alone, together: new forms of commensality. In Proceedings of the 17th ACM international conference on Supporting group work - GROUP '12, 103. Sanibel Island, Florida, USA: ACM Press. <https://doi.org/10.1145/2389176.2389192>.
- Harboe, Gunnar, Crysta J. Metcalf, Frank Bentley, Joe Tullio, Noel Massey, and Guy Romano. 2008. Ambient social tv: drawing people into a shared experience. In Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08, 1. Florence, Italy: ACM Press. <https://doi.org/10.1145/1357054.1357056>.
- Hassenzahl, Marc, Stephanie Heidecker, Kai Eckoldt, Sarah Diefenbach, and Uwe Hillmann. 2012. All You Need is Love: Current Strategies of Mediating Intimate Relationships through Technology. *ACM Transactions on Computer-Human Interaction* 19: 1–19. <https://doi.org/10.1145/2395131.2395137>.
- IJsselsteijn, Wijnand, Joy van Baren, and Froukje van Lanen. 2003. Staying in Touch Social Presence and Connectedness through Synchronous and Asynchronous Communication Media. In *Human-Computer Interaction: Theory and Practice*, II:5. Crete, Greece.
- Kirk, David S., and Abigail Sellen. 2010. On human remains: Values and practice in the home archiving of cherished objects. *ACM Transactions on Computer-Human Interaction* 17: 1–43. <https://doi.org/10.1145/1806923.1806924>.
- Kusenbach, Margarethe. 2003. Street Phenomenology: The Go-Along as Ethnographic Research Tool. *Ethnography* 4: 455–485. <https://doi.org/10.1177/146613810343007>.
- Kuutti, Kari, and Liam J. Bannon. 2014. The turn to practice in HCI: towards a research agenda. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 3543–3552. Toronto Ontario Canada: ACM. <https://doi.org/10.1145/2556288.2557111>.
- Kuwabara, K., T. Watanabe, T. Ohguro, Y. Itoh, and Y. Maeda. 2002. Connectedness oriented communication: fostering a sense of connectedness to augment social relationships. In Proceedings 2002 Symposium on Applications and the Internet (SAINT 2002), 186–193. Nara, Japan: IEEE Comput. Soc. <https://doi.org/10.1109/SAINT.2002.994476>.
- Liaqat, Amna, Bennett Axtell, and Cosmin Munteanu. 2021. Participatory Design for Intergenerational Culture Exchange in Immigrant Families: How Collaborative Narration and Creation Fosters Democratic Engagement. *Proceedings of the ACM on Human-Computer Interaction* 5: 1–40. <https://doi.org/10.1145/3449172>.
- Liechti, Olivier, and Tadao Ichikawa. 1999. A Digital Photography Framework Supporting Social Interaction and Affective Awareness. In *Handheld and Ubiquitous Computing*, ed. Hans-W. Gellersen, 1707:186–192. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-48157-5_18.

- Light, Ann, and Daniela Petrelli. 2014. The rhythm of Christmas: temporality, ICT use and design for the idiosyncrasies of a major festival. In Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design, 159–167. OzCHI '14. Sydney, New South Wales, Australia: Association for Computing Machinery. <https://doi.org/10.1145/2686612.2686636>.
- Lottridge, Danielle, Nicolas Masson, and Wendy Mackay. 2009. Sharing empty moments: design for remote couples. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2329–2338. Boston MA USA: ACM. <https://doi.org/10.1145/1518701.1519058>.
- Lynggaard, Aviaja Borup, M. G. Petersen, R. Gude, and M. Mortensen. 2010. Home awareness: connecting people sensuously to places. In Proceedings of the 8th ACM Conference on Designing Interactive Systems - DIS '10, 416. Aarhus, Denmark: ACM Press. <https://doi.org/10.1145/1858171.1858251>.
- Min, Hyeryung Christine, and Tek-Jin Nam. 2014. Biosignal sharing for affective connectedness. In CHI '14 Extended Abstracts on Human Factors in Computing Systems, 2191–2196. Toronto Ontario Canada: ACM. <https://doi.org/10.1145/2559206.2581345>.
- Morris, Margaret E., Douglas M. Carmean, Artem Minyaylov, and Luis Ceze. 2017. Augmenting Interpersonal Communication through Connected Lighting. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, 1924–1931. Denver Colorado USA: ACM. <https://doi.org/10.1145/3027063.3053141>.
- Mueller, Florian “Floyd,” Frank Vetere, Martin R. Gibbs, Jesper Kjeldskov, Sonja Pedell, and Steve Howard. 2005. Hug over a distance. In CHI '05 Extended Abstracts on Human Factors in Computing Systems, 1673–1676. Portland OR USA: ACM. <https://doi.org/10.1145/1056808.1056994>.
- Mynatt, Elizabeth D., Jim Rowan, Sarah Craighill, and Annie Jacobs. 2001. Digital family portraits: supporting peace of mind for extended family members. In Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '01, 333–340. Seattle, Washington, United States: ACM Press. <https://doi.org/10.1145/365024.365126>.
- Odom, William, James Pierce, Erik Stolterman, and Eli Blevis. 2009. Understanding why we preserve some things and discard others in the context of interaction design. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1053–1062. Boston MA USA: ACM. <https://doi.org/10.1145/1518701.1518862>.
- Odom, William, John Zimmerman, and Jodi Forlizzi. 2010. Designing for dynamic family structures: divorced families and interactive systems: 10.
- Rettie, Ruth, and RM Rettie. 2003. Connectedness, Awareness and Social Presence. In , 7.
- Romero, Natalia, Panos Markopoulos, Joy van Baren, Boris de Ruyter, Wijnand IJsselsteijn, and Babak Farshchian. 2007. Connecting the family with awareness systems. *Personal and Ubiquitous Computing* 11: 299–312. <https://doi.org/10.1007/s00779-006-0089-0>.
- Short, J., E. Williams, and B. Christie. 1976. *The social psychology of telecommunications*. London: Wiley.
- Swan, Laurel, Alex S. Taylor, and Richard Harper. 2008. Making place for clutter and other ideas of home. *ACM Transactions on Computer-Human Interaction* 15: 1–24. <https://doi.org/10.1145/1375761.1375764>.
- Taylor, Alex S., and Laurel Swan. 2005. Artful systems in the home. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 641–650. Portland Oregon USA: ACM. <https://doi.org/10.1145/1054972.1055060>.

- Tibau, Javier, Michael Stewart, Steve Harrison, and Deborah Tatar. 2019. FamilySong: Designing to Enable Music for Connection and Culture in Internationally Distributed Families. In Proceedings of the 2019 on Designing Interactive Systems Conference, 785–798. San Diego CA USA: ACM. <https://doi.org/10.1145/3322276.3322279>.
- Van Bel, Daniel T., Wijnand A. IJsselsteijn, and Yvonne A.W. de Kort. 2008. Interpersonal connectedness: conceptualization and directions for a measurement instrument. In CHI '08 Extended Abstracts on Human Factors in Computing Systems, 3129–3134. Florence Italy: ACM. <https://doi.org/10.1145/1358628.1358819>.
- Wakkary, Ron, and Leah Maestri. 2007. The resourcefulness of everyday design. In Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition - C&C '07, 163. Washington, DC, USA: ACM Press. <https://doi.org/10.1145/1254960.1254984>.
- Wallbaum, Torben, Andrii Matviienko, Swamy Ananthanarayan, Thomas Olsson, Wilko Heuten, and Susanne C.J. Boll. 2018. Supporting Communication between Grandparents and Grandchildren through Tangible Storytelling Systems. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 1–12. Montreal QC Canada: ACM. <https://doi.org/10.1145/3173574.3174124>.
- Wong-Villacres, Marisol, and Shaowen Bardzell. 2011. Technology-mediated parent-child intimacy: designing for ecuadorian families separated by migration: 6.
- Wyche, Susan P, and Marshini Chetty. 2013. “I want to imagine how that place looks”: designing technologies to support connectivity between africans living abroad and home. In , 10.
- Ylirisku, Salu, Antti Jylhä, Anu Lehtiö, Imtiaj Ahmed, Craig Stewart, Abigail Sellen, Richard Harper, and Giulio Jacucci. 2016. Designing for Active Place Presence at Home: The Hole in Space Design Experiment. In Proceedings of the 2016 ACM Conference on Designing Interactive Systems, 1341–1352. Brisbane QLD Australia: ACM. <https://doi.org/10.1145/2901790.2901884>.
- Ylirisku, Salu, Siân Lindley, Giulio Jacucci, Richard Banks, Craig Stewart, Abigail Sellen, Richard Harper, and Tim Regan. 2013. Designing web-connected physical artefacts for the “aesthetic” of the home. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 909–918. Paris France: ACM. <https://doi.org/10.1145/2470654.2466117>.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

Xinchen Yu, Afra Mashhadi, Jeremy Boy, Rene Clausen Nielsen and Lingzi Hong (2022): Causal Impact Model to Evaluate the Diffusion Effect of Social Media Campaigns. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centered Computing on the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_ep10

Causal Impact Model to Evaluate the Diffusion Effect of Social Media Campaigns

Xinchen Yu¹, Afra Mashhadi², Jeremy Boy³, Rene Clausen Nielsen⁴, Lingzi Hong¹

¹University of North Texas

²University of Washington Bothell

³UNDP Accelerator Labs

⁴UN Global Pulse
xinchenyu@my.unt.edu

mashhadi@uw.edu

jeremy.boy@undp.org

rene@unglobalpulse.org

lingzi.hong@unt.edu

Abstract. Organized information campaigns on social media platforms have influence on collective opinions through processes such as social influence and majority opinion formation. Evaluating the effect of such campaigns has become a critical question. We proposed a method by first characterize user engagement and the semantics in public discussions with social media data, then apply a causal impact analysis to measure the effect. We conducted a case study to examine the effect of the 16 Days Campaign (a campaign organized by UN Women) through changes in public discussions of the MeToo, which is a related topic the campaign was aimed to impact. Results showed there were significantly more discussions in MeToo after the launch of the campaign. Hashtags on 16Days topics were used more and by more people. The proposed methods evaluate the direct and indirect diffusion effect of a campaign by quantifying the difference had the campaign not taken place based on social media data. The method enables to evaluate the overall outcome of collaborative work in a social media campaign.

Introduction

Social media campaigns refer to campaigns sustained by coordinated efforts to achieve specific goals with information spreading on social media platforms. Participants or volunteers are guided to post certain content under the agenda of

the campaign by using strategies such as spreading certain propaganda or diverting public attentions from important issues to shape collective opinions online (Bradshaw and Howard, 2018). Collective opinions may lead to individual behavioral changes, such as voting preferences (Aral and Eckles, 2019), or societal changes in policy making and legislation (Fileborn and Loney-Howes, 2019; Freischlag and Faria, 2018). How to evaluate the collaborative effort and the impact of social media campaigns has become a critical question. Campaigns oftentimes implement different strategies, for example, promoting through influencers or coalition with trending topics. The campaign managers hope to know whether the campaign achieved the expected outcomes by attracting public attentions or influencing people. Evaluating the effect enables to compare and identify effective campaigns. It may also help defense the impact of malicious campaigns that spread conspiracy theories or misinformation (Badawy et al., 2019).

Different methods have been implemented to investigate the effect of social media campaigns. Most of the evaluations are based on campaign participants, thus, to survey participants before and after campaigns (Thompson et al., 2020) or to compare the differences between campaign participants and controlled groups (Breza et al., 2021). However, many social media campaigns do not have either the controlled group setting or identifiable participants, not to mention that the collection of individual data such as survey responses or digital traces is difficult and costly (Aral and Eckles, 2019). In addition, social media campaigns are found to have spillover effect in many ways (Lee et al., 2018; Dincelli et al., 2016). Such effect cannot be measured with only data from campaign participants. Other studies examine outcomes of social media campaigns by analyzing the related tweets after certain events. However, tweets on the examined topics may exist before the campaign, the results may overestimate the effect of the campaign. The results are mostly descriptive with visualizations showing the temporal trend after the events. For example, Samuel et al. (2020) analyzed the emotional consequences of the public after the reopening policy during the COVID-19. Quantitative evaluation of the effect, thus, to measure the degrees of change had the event not happened is few. Recently, several studies proposed to statistically analyze the relation between social media campaigns and certain voting (Aral and Eckles, 2019) or protesting behaviors (Ahmed et al., 2017) with causal inference analysis. Few have applied the causal impact model to the evaluation of collective opinions impacted by social media campaigns. We apply the causal impact model to analyze the relative effect of a social media campaign, the 16Days campaign, on collective opinions in MeToo.

The study presented here has two main contributions. First, we design a set of metrics to characterize users' participation and the trending topics in collective opinions with social media data. The metrics can capture the diffusion effect of social media campaigns with data at the aggregated level, which doesn't require

the identification of campaign participants or the use of individual level data. Second, we present a causal inference method for the quantitative evaluation of the diffusion effect of a social media campaign, which measures the difference between the observed values and the counterfactual values had the social media campaign not taken place. The causal impact model enables to measure the accumulated impact during a time range.

We applied the method on a case study and examined changes in the MeToo discussions after the launch of a social media campaign, the 16Days, which used hashtags such as #HearMeToo to raise discussions in MeToo and geared towards the awareness of women's rights and empowerment. Results showed that the 16Days campaign may have led to increased engagement in MeToo and more discussions on women empowerment topics, illustrating how a social media campaign may indeed influence public discussions.

The rest of this article is organized as follows: In Section 2, we review the related work on information diffusion on social networks and methods to evaluate the effect of social media campaigns. Section 3 presents on research methodology including an overview of the causal impact model, and the data we used for the case study. Section 4 presents our analysis and findings. Section 5 discusses the method and implications. Section 6 concludes this article and the limitations.

Related Work

Organized social media campaigns have been found to be implemented in many countries, which use strategies such as promoting certain content to cause it a trend, shaping discussions through comments, attacking opponents, and diverting conversations from important issues. Some of the campaigns could become a potential threat to the democratic society (Bradshaw and Howard, 2018).

Diffusion Effect in Social Networks

Several studies have investigated the diffusion of campaign information on social media. For example, Badawy et al. (2019) characterized the diffusion paths of messages from Russian Internet Research Agency's information campaigns during the 2016 US presidential election; Ferrara et al. (2020) analyzed the diffusion of automated misinformation in different user groups during the 2020 US presidential election. Recent studies have also proposed methods to identify the initial set of nodes (Smith et al., 2018) or interaction of nodes (Myers and Leskovec, 2012) that can maximize the diffusion of information in social networks.

However, the impact of social media campaigns is not only limited to the users involved in the diffusion of campaign information. A line of research investigates mechanisms how campaign information influence public opinions on social

media. The diffusion of campaign information can be amplified under the effects such as social affordance (Lee et al., 2018), homophily (Dincelli et al., 2016), and social contagion (Coviello et al., 2014). Homophily describes nodes sharing similar characteristics tend to act out similarly (Dincelli et al., 2016). Social contagion is used to characterize mimicry practice, such as the adoption or formation of opinions from social contacts due to network diffusion (Christakis and Fowler, 2013). The diffusion between agents is found to be moderated by many factors, for example, the sociocultural distance (DellaPosta et al., 2015; Yu et al., 2020 (a); Yu et al., 2020 (b)) and the interpreted meaning of the practice based on the distribution of all practices in the population (Goldberg and Stein, 2018). The initial exposure increases probability of diffusion, but the probability can be suppressed when the exposure comes to saturation (Hodas and Lerman, 2014). Yoo et al. (2019) built a self-exciting point process model to examine how the diffusion of information might be influenced by the parallel diffusion of similar content. Results showed that the diffusion effect can be inhibited or amplified depending on the network structure. These studies show the impact of a social media campaign is not only through the direct diffusion of campaign information but can be more intricate and complex. Therefore, the evaluation of campaign effect is partial if only campaign participants are considered, it requires an examination of the systematic outcome.

Evaluating the Effect of Campaigns

A lot of recent studies have paid attention to evaluating the effect of social media campaigns with a focus on campaign participants. These studies were conducted by comparing participants before and after the campaign (Thompson et al., 2020), people who are in campaign regions and are not (Buller et al., 2021), or people in regions with campaign activities of high and low intensity (Breza et al., 2021). For example, Thompson et al. (2020) evaluated the effect of a social media campaign on changing the mental health stigma in student participants. Breza et al. (2021) used behavioral data collected from Facebook to investigate how a social media campaign affected the mobility behaviors of people during COVID-19. The evaluation relies on the implementation of comparative experiments with controlled groups or is based on the survey data or digital traces from individuals. Such data may not be readily available in many cases. Most of the campaigns are not designed to be launched in certain geographical areas. It is also difficult to identify campaign participants and collect behavioral data from individuals (Aral and Eckles, 2019). In social media campaigns, people who are not campaign participants could be affected through relationships and the diffusion of information (Dincelli et al., 2016; Coviello et al., 2014). These methods cannot measure the spillover effect.

Aral and Eckles (2019) proposed to use causal inference to evaluate how social media manipulation may have influenced the US presidential election outcome.

Causal statistical analysis can be applied to analyze public opinion and behavior change across individuals and subpopulations by measuring deviations from expected behaviors due to manipulative factors. In this study, we propose to apply the causal impact method to evaluate the outcomes on collective opinions with social media data.

Social Media Data to Study Public Opinions

A lot of public discussions are happening on social media platforms. Social media data, therefore, become an important source of information to understand public opinions. Several studies have used social media data to analyze the consequences of events such as protests (De Choudhury et al., 2016, Ahmed et al., 2017; He et al., 2015), implementation of new policies (Samuel et al., 2020), public health incidents (Gaspar et al., 2016), and social media campaigns (Badawy et al., 2019). Samuel et al. (2020) focused on characterizing the public sentiment trends for the policy decision in COVID-19 pandemic. Hong et al. (2016) analyzed the trend of different topics after the Ferguson unrests. These studies generate descriptive results about the sentiment and topics of discussions and interpret the relations of the results to the events. Some studies used statistical analysis, for example, De Choudhury et al. (2016) unpacked the relations of thoughts, opinions, and sentiments in the discussions of the “Black Lives Matter” on Twitter and perspectives of offline protests. Ahmed et al. (2017) discovered a high resemblance in emotions of anger and anxiety between online discussions and offline campaigns. The studies are often used to modeling collective opinions echoing to offline events. Few have applied the causal impact method to analyze the relative effect on public opinions impacted by social media campaigns.

Research Methodology

This study aims to measure the impact of a social media campaign on the collective discussions on a related theme, identifying if the campaign leads to more public attention on the campaign agenda. On social media platforms such as Twitter, discussions on certain topics are symbolized by a set of hashtags. A campaign can interact with users using certain hashtags and expose the campaign claims to people sharing similar interests and impose influence. The diffusion of campaign information may lead to increased participation on certain topics and change public opinions through the process of social influence (campaign claims spread to people through social contacts) and homophily (users with similar interest show conformity in spreading information of the campaign) (Dincelli et al., 2016). Further, the campaign may bring public attention on specific claims and bring more discussions related to the claims (Edwards and Marullo, 1995). The dynamics of the diffusion process can be complex under the effects of

confounding factors (Monsted et al., 2017). We focus on the method to evaluate the outcomes of diffusion, considering that such interactions may lead to participation and content changes in collective opinions.

We conducted a case study on the effect the 16 Days of Activism (16Days) on public discussions of the MeToo. The 16Days is an annual international campaign sponsored by United Nations General Secretary and organized by UN Women that kicks off on 25 November 2018, the International Day for the Elimination of Violence against Women, and runs until 10 December, Human Rights Day. In 2018, the 16 Days campaign was designed to be associated with the MeToo movement online using hashtags such as #HearMeToo. The online campaign may trigger contagious ideas or behaviors in the discussions of MeToo. We aimed to evaluate the diffusion effect of the campaign and monitored changes in the MeToo. The campaign may have influenced public discussions in other topics that were not part of the MeToo. However, it is impossible to monitor all discussions on social media platforms.

Data Collection

Since the base of the campaign is in the US and most of the campaign conversations are in English, we used English tweets for the evaluation. With the help of UN Women specialists, hashtags of the MeToo movements and the 16Days campaign were identified (Table I). We used Crimson Hexagon ForSight, now Brandwatch Consumer Research (Crimson Hexagon, 2019), to collect two types of data: i) tweeting activities in the MeToo that exclude tweets with 16Days campaign hashtags. We collected time series data such as daily volumes of all tweets and unique tweets; and ii) sampled MeToo and 16Days tweets that are used for topic analysis. In total, there were 32,249,394 tweets posted in the MeToo movement from January 1, 2016, to January 15, 2019. We collected 313,451 sampled tweets, 151,398 of which are tweets of the 16Days campaign from 66,619 users in a 62 days' time span, 162,053 are tweets of the MeToo movement from 110,522 users in a 1,111 days' time span.

Table I. Hashtags and sampled tweets of the MeToo movement and the 16Days campaign.

	Hashtags	#Tweets	#Users	#Days
MeToo movement	#MeToo, #WithYou, #WeToo, #TimesUp, #TimeisNow	162,053	110,522	1,111
16Days campaign	#GBVteachIn, #GBV, #endGBV, #16daysofactivism, #endVAW, #endVAWG, #womensrights, #humanrights, #ILOendGBV, #StopGBVatWork, #RatifyILO190, #16days, #orangetheworld, #16DaysCampaign, #HearMeToo	151,398	66,619	62

Characterizing User Participation and Trending Topics in Time Series

We propose four metrics to measure the scale of user participation in MeToo: 1) total volumes of the posts, which reflects the scale of public discussions on the topic and partially reflected the occupation of public attention; 2) total volumes of unique posts, which measures the amount of new content created; 3) number of unique links in posts, which measures the diversity of exogenous information sources, such as news websites and videos; and 4) number of unique hashtags used, which reflects the diversity of information generated from the social media platform. All are computed in a certain temporal granularity. In the case study, we compute daily values.

Topical changes are measured based on the use of hashtags. Daily frequency is computed as number of posts with the hashtag. Considering the total volumes of tweets on MeToo could vary by day, daily proportion is used, which measures the proportion of tweets containing a certain hashtag to the total volume of MeToo posts. Active participants of a topic may advocate similar content for hundreds or even thousands of times (Shao et al., 2018). Therefore, we introduce daily weight in participants, which is the average ratio of a hashtag used by all users in MeToo. The metric is to capture the weight of a topic in the MeToo discussion. To further measure whether the hashtags with significant changes are relevant to the campaign, we compute the semantic similarity of the hashtag to the MeToo hashtag and campaign hashtag. The semantic embedding of these hashtags is learned by the global vectors for word representation (GloVe) algorithm. The GloVe embedding algorithm incorporates global statistics, i.e., word co-occurrence probabilities, and local context information, i.e., the neighboring words in a window, to obtain vector representations for words (Pennington et al., 2014). The similarity of two hashtags is measured by cosine similarity of the vector representations of hashtags.

Causal Impact Analysis Model

We conducted the causal impact analysis by applying the state-space model, which predicts the counterfactual time series indicating the status had no intervention taken place based on control groups (Brodersen et al., 2015). The model is chosen over other causal inference models as the effect is evaluated based on time series data. The model has two components: (1) the observed values capture the trend of the states and the disturbance in the observation. (2) the state is assumed to be a latent process with variances due to noises. The hidden or latent process is assumed to be a Markov process. The states generate observations, and these observations are only dependent on the states. The state is composed of local linear trend, seasonality, and contemporaneous covariates with static coefficients.

A structural model can be inferred based on the observed values in control groups (Brodersen et al., 2015). The observed values are related to the states, while the states are inferred from seasonal patterns, previous states, and potential noises. The model considers noisy signals, for example, users may discuss other issues using movement hashtags and contribute to the total volume of movement discussion. The model can be used to predict, given the unchanged dynamics of states and the existence of disturbance, what the observations could be. We can thus build a structural model based on historical observations and compare the observed values with the predicted values had the intervention not taken place.

The differences between the observed time series and the counterfactual time series reflect the impact of an intervention event. Time-series data usually have fluctuations. It is difficult to determine the impact level at a discrete time point. We measure the impact by accumulating all the differences in a time range and compute the relative effect in total. To make sure the effect size is comparable over time series of different measurements, we compute the relative effect as a percentage of summed differences to the summed counterfactual time series values. The causal impact analysis will be applied to the following measurements using a Python package Causal Impact (Fuks, 2020).

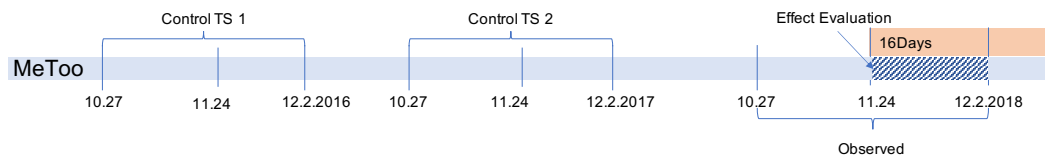


Figure 1. Control groups and effect evaluation time window in the timeline of MeToo.

Effect Evaluation

The 16Days campaign used the hashtag #HearMeToo to encourage people participate in the MeToo movement from Nov 24, 2018. The 16Days campaign can be seen as an event that cause variances in the observed time series of MeToo. The effect was evaluated by comparing the observed time series of MeToo with the synthetic time series (counterfactual), which were inferred based on historical data (control groups) assuming no interventions had happened.

The compared time series was set as 1 week from the time of the intervention with a granularity of one day. Correspondingly, the counterfactual data was about the same week that consisted of seven daily values. To infer the counterfactual time series, we set control groups with three components: i) the observed time series in 4 weeks before the intervention, which is Oct 27, 2018 to Nov 24, 2018 (part of Observed in Figure 1); ii) the observed time series one year ago, with 4 weeks before the intervention date and 1 week after, which is Oct 27, 2017 to Dec 2, 2017 (Control TS2 in Figure 1); iii) the observed time series two years ago, with the five weeks the same as ii), i.e. Oct 27, 2016 to Dec 2, 2016 (Control TS1

in Figure 1). These parameters are applied to evaluate time series behaviors in the scale of user participation and discourses.

In terms of topic analysis, we identified the hashtags that had significant changes due to the campaign intervention and computed the similarity of these hashtags to #HearMeToo and #MeToo, which are separately key hashtags of the 16Days campaign and the MeToo. The GloVe embeddings of hashtags were computed with the tweets of MeToo and 16Days in Control TS1, Control TS2 and Observed time ranges. After removing the repetitive tweets, we pre-processed the tweets by removing links, special characters, punctuation symbols, and entities that start with "@", transforming all tokens to lower cases and stemming words. After pre-processing, we got the vector representation of all the hashtags with a length of 50 and computed the cosine similarity of hashtags.

Finding

Figure 2 shows the contrast between counterfactual values and the observed volumes of tweets from October 27 to Dec 2 of 2018 for all tweets that include retweets (a) and unique tweets (b). Both figures show that the observed volumes are significantly larger than the model predicted values based on historical time-series (counterfactual). The increasing trend reached to the max in the first three days since the launch of the campaign. After that, the difference between observed values and the counterfactual values becomes smaller, which shows the effect of the campaign weakens.

Table II. Relative effect of causal impact analysis.

	In one week		In two weeks	
	relative effect	95% interval	relative effect	95% interval
Daily volume	38.44%	(7.71%, 68.62%)*	28.93%	(7.89%,50.18%)*
Daily volume of original tweets	34.02%	(11.32%, 56.15%)*	37.67%	(21.41%,55.62%)*
Proportion of original tweets	-4.85%	(-16.82%, 8.0%)	3.45%	(-6.74%,13.4%)
Number of unique links	1.8%	(-8.74%, 12.57%)	-10.98%	(-18.36%,-3.38%)*
Number of unique hashtags	13.16%	(-2.34%, 28.21%).	-6.02%	(-17.46%,5.35%)

Significance levels: 0 ‘***’; 0.001 ‘**’; 0.01 ‘*’; 0.05 ‘.’; 0.1 ‘.’

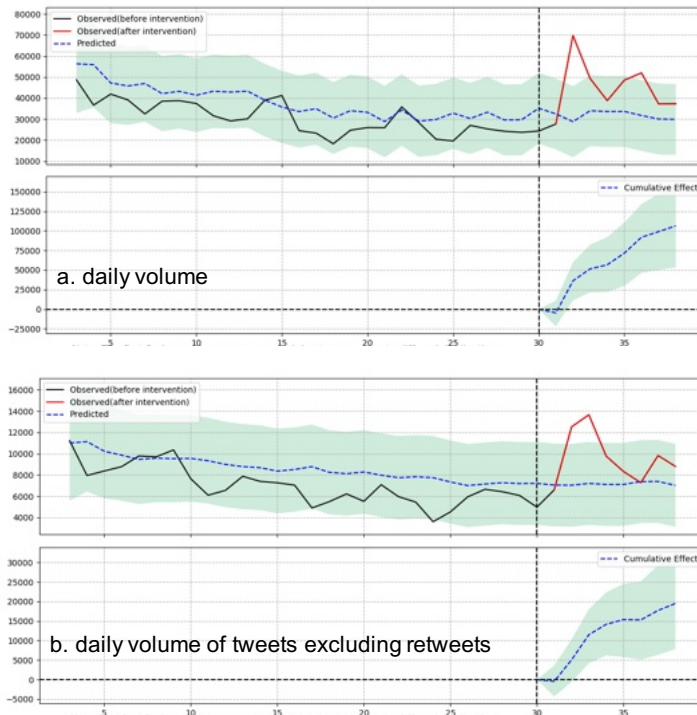


Figure 2. Time series of predicted and observed behaviors.

Table II shows the statistics of causal impact analysis. We found increased discussions in MeToo after the launch of the 16Days campaign, with more unique posts created by the participants and more retweeting. The observed volume of MeToo tweets reached to the maximum in the third and fourth day after the campaign was launched, then the effect was weakened afterwards. It is possible that the diffusion of campaign related information reached to the point of saturation, which inhibited the creation and diffusion of information (Hodas and Lerman, 2014).

Associated with the increasing volume, the unique external links contained in MeToo tweets did not change significantly but the unique hashtags increased. It reflects the increase of posts may come from the introducing of a new set of topics on Twitter rather than from the sharing of information from external sources. The causal impact model has considered potential noises and variances caused by random factors and attribute the effect of change to something happened at a certain time. In the case study, we set it as the day when the 16Days campaign was launched. The analysis of unique links and hashtags shows the increased participation of MeToo is probably not from other trending events, as the unique links indicating external sources do not change significantly in the first week. Rather, the increasing trend is more associated with information initiated from the platform, which is reflected from the more hashtags and more original content created in MeToo. To further reveal how the introduction of new

topics is related to the 16Days campaign, we conducted semantic analysis for online discussions.

Table III. Relative effect and the 95% interval of relative effect of hashtag usage. Top 10 hashtags with the largest relative effect are shown.

Hashtags	Weights in participants	Proportion of use	Frequency of use	Similarity	
				#MeToo	#HearMeToo
#freedom	25.11 (22.4, 27.81)***	45.93 (44.13, 47.61)***	39.14 (37.65, 40.89)***	0.44	0.32
#witchhunt	11.73 (10.6, 12.92)***	-0.16 (-1.35, 0.97)	-0.24 (-1.38, 1.02)	0.08	0.08
#womensmarch	5.55 (4.5, 6.55)***	3.98 (3.26, 4.66)***	3.59 (2.86, 4.3)***	0.35	0.39
#democracy	4.28 (2.41, 5.93)***	26.35 (24.91, 27.95)***	22.61 (21.05, 24.32)***	0.33	0.20
#survivor	3.22 (2.31, 4.13)***	1.3 (0.21, 2.31)**	1.4 (0.34, 2.38)***	0.47	0.37
#justice	2.95 (1.42, 4.69)***	25.97 (24.34, 27.69)***	24.07 (22.45, 25.76)***	0.48	0.27
#metoomvmt	2.24 (1.26, 3.18)***	1.36 (0.4, 2.42)***	1.56 (0.59, 2.54)***	0.31	0.33
#brexit	1.93 (0.92, 3.06)***	0.28 (-0.83, 1.36)	0.31 (-0.77, 1.36)	0.15	0.27
#nomore	1.71 (0.29, 3.22)**	0.12 (-1.12, 1.38)	0.1 (-1.17, 1.34)	0.25	0.21
#womenempowerment	1.63 (0.95, 2.25)***	0.28 (-0.6, 1.22)	0.2 (-0.77, 1.1)	0.30	0.39

Significance levels: 0 '***'; 0.001 '**'; 0.01 '*'; 0.05 '.'; 0.1 ' '.

For each hashtag that has been used in MeToo posts from 4 weeks before Nov 24 and 1 week after, we counted its daily frequency, daily proportion, and daily weight by users in TS1, TS2, and the observed time range. The values were set to zero if no hashtag was detected to be used in a day. We conducted a causal impact analysis with the time series of the three features for all hashtags. Table 3 lists the top 10 hashtags that show significant positive effect in daily weight by users in one week after the launch of the 16Days campaign, as we focused on topics that emerged due to the campaign. We also included the relative effect in terms of daily frequency and daily proportion as cross-references. We didn't present the 16Days hashtags here as they do not exist in previous MeToo tweets, thus all have significantly positive relative effect after the campaign.

There is a significant increase in the use of #freedom, which had a relative effect of 2511% in daily weight in users. The effect on daily frequency and daily

proportion is even larger. #Freedom was used in different contexts, for example, to call for a freedom culture from traditions – “... go out against #patriarchy, #misogyny and other abuse of women...#freedom #equality #secular”, or to talk about the freedom of sex victims of human trafficking – “Let's end slavery... #freedom #slavery #escort #childabuse #metoo #stopthedemand...”. The hashtag has a relatively higher similarity with #MeToo but is also seen in many cases that have gender perspectives. For example, “@XXX on #MeToo #GenderEquity #Freedom #quotes Gender justice has never been only about women, but about modernity, an expansion of democracy, ...”

Of the 10 hashtags, #womensmarch, #metoomvmt, and #womenempowerment have higher similarity with #HearMeToo compared to #MeToo. There is an estimated increase of 555% in #womensmarch, 224% in #metoomvmt, and 163% in #womenempowerment in terms of the daily weight in users. Both #womensmarch and #womenempowerment have an emphasis on the gender perspectives, which have overlaps with the claims of the 16Days campaign. The hashtag #womensmarch originated from a protest anti-women and offensive statements by politicians in 2017. The movement is mainly led by women and has a critical goal to end gender-based violence. Women's March is scheduled every year around January 20th from 2017. The variances cannot be a post-event effect as there are ten months' lag. The significantly increasing weight of #womensmarch could be partially attributed to the campaign. Hashtag #womenempowerment is mainly used in posts on gender equality. It calls for the inclusion of women in the decision-making process in economy and politics through “education, raising awareness, literacy and training” (Bayeh, 2016). The trending of these hashtags indicates more discussions on the gender-based claims and gender equality, which is what the 16Days advocates. It implies that the diffusion of campaign information may have lead participants in the MeToo movement to have increased attention on topics advocated by the campaign. Meanwhile, there is increased use of hashtags such as #freedom, #survivor, #justice, and #nomore, which although have higher similarity with #MeToo but are used in the 16Days claims that calls for social justice to sex abuse survivors. Besides, we observed the increased weight of #witchhunt, which is not quite related to the MeToo movement or the 16Days campaign, and #democracy and #brexit, which are mainly used in political discussions. The findings show that not all the increased discussions were caused by the diffusion effect of the campaign.

Discussion

To summarize the findings, we found that novel information was generated and circulated in the MeToo after the launch of the 16Days campaign even after removing the tweets that are directly brought by participants of the campaign. The increasing trend is associated with significantly more hashtags rather than the

number of external links, which could indicate the effect mainly come from online events. Further, the semantic analysis shows the weight of some topics (hashtags) has significantly increased, of which there are #womensmarch, #womenempowerment and #metoomvmt that have a higher association with the 16Days campaign. The results illustrate other external factors than the campaign, may together exert confounding impact on MeToo discussions.

In the case study, we evaluated the impact with public opinions on MeToo due to the relation between the campaign and MeToo discussions. To apply the method on other effect evaluation cases, there are several premises to make sure the results are relatable to the intervention of the event. The monitored activities should be an expected outcome brought by the campaign. The method helps to find out how much the monitored activities could be affected. Social media discussions could be subjected to external factors such as posts of influential figures or news. If there are identifiable factors, the effect of these factors should be considered in the interpretation of results.

The choice of controlled time series for training the model to generate counterfactual values is important. In the presented case, we chose controlled time series from two years. The choice is partially due to the lack of data before 2016 for model training, as there are too few MeToo discussions before 2016. Each year we used activities in four weeks as previous states to infer the following week at the same time of the year. This allows to capture seasonable patterns. A longer time range for training may include large variations caused by significant events in previous years, while a shorter time range doesn't include enough data points to train the model. In other cases, effect evaluation analysis may consider similar reasons for the choice of time series to build the model.

We combined different perspectives of analysis to infer the impact of the campaign. It is highly probable the outcome doesn't attribute to one single factor. In fact, due to the existence of confounding factors, the identified relative effect is a maximum boundary for reference. We may use topic analysis to further investigate the relations. The method provides a complementary analysis that focused on the deviation from normal patterns had the event not taken place. It may not apply to new activities caused by the event, for example, the use of hashtags created by the campaign was not used before. The method is especially valuable to evaluate the differences from normal patterns caused by a campaign.

Finally, in the case study we mainly analyzed the use of hashtags. Combing with other types of content analysis, we may generate metrics that characterize different dimensions of public opinions. For example, we may generate time series of sentiment polarity to investigate if the campaign has significant emotional consequences in the public. By using automatic classification algorithms, we may identify tweets that belong to the misinformation categories. It will enable to identify whether a certain event or campaign led to significant increase or decrease of misinformation spreading. In this case, we mainly

explored the measurements that are valuable to evaluate the effect of the 16Days campaign.

Implications

Social media campaigns are collaborative work by campaign participants but could have impact on users who are not directly engaged in the campaign through diffusion effects or public attentions. Therefore, effect evaluation of campaigns should consider the overall outcomes rather than focus on the effects on participants. The proposed method enables to quantitatively evaluate the differences that can be induced by an event. Different from previous studies that focus on interpreting descriptive statistics of the outcomes after an event, the method considers the activities that could have existed due to previous states and seasonal patterns. It provides statistics indicating if the relative effect is significant due to the events or if the variances are generated by random noises.

The results presented have practical implications for campaign managers and activist organizations. First, our results echo the findings that an organization-driven campaign could be highly influential and impactful in shaping online discussions. In the case of MeToo, a movement that is often associated with celebrities, our results show that a timely and unified intervention by UN Women helped to spread the MeToo movement and broadening the movement through diversifying. Second such diversification could also lead to a more diverse set of audience and reach well beyond the celebrity focus. The method proposed in this study can be applied to the effect evaluation of other social media campaigns and allow for dynamic monitoring of the intervention by the campaign managers.

Conclusion

This study investigates the diffusion effect of a campaign on public discussions. We have proposed measurements to characterize movement participation from the perspectives of user engagement and discourses with social media data. We have applied a causal impact analysis to measure the relative effect due to the intervention of the campaign. The method is helpful to find the maximum boundary of the relative effect on collective opinions. Combining semantic analysis, it further reveals how the campaign influenced the discussions and other potential factors that might have affect the discussion.

This study is an initial exploration with several limitations. First, we have tried to capture the evolving of the collective opinions from different perspectives but there are other ways to characterize the participation and topics we did not include, for example, the number of individual participants or mentioned users. The analysis focused on the metrics that we could obtain through the summary statistics from the API (Brandwatch). That's the case for data-driven studies that

we can only develop methods and reveal insights based on the available data. However, it is notable that the causal impact analysis method is applicable to different time series data at different temporal granularity for evaluating the effect. Given more data, we will be able to present a more comprehensive evaluation. Second, the causal impact model we used has limitations. It is used to measure the differences had something not happened in a certain time point. As we mentioned, public opinions could be potentially affected by other factors that happen near the time point. It is possible there were chain reactions, for example, the increased public attention led politicians to involve in the discussions and bring new topics. However, the current state-space model has limitations that it cannot model the unidentified external factors or the effect of a series of external factors. Third, the case study presented here investigates the campaign in 2018, which was designed to be related to the MeToo movement. It would be a future work to obtain data from other years to compare the effect and identify effective strategies.

Acknowledgments

The authors would like to thank UN Global Pulse Lab and UNWomen for their time and support to shape this work.

References

- Ahmed, S., Jaidka, K. & Cho, J. (2017), 'Tweeting india's nirbhaya protest: a study of emotional dynamics in an online social movement', *Social Movement Studies* 16(4), 447–465.
- Aral, S. & Eckles, D. (2019), 'Protecting elections from social media manipulation', *Science* 365(6456), 858–861.
- Badawy, A., Addawood, A., Lerman, K. & Ferrara, E. (2019), 'Characterizing the 2016 russian ira influence campaign', *Social Network Analysis and Mining* 9(1), 1–11.
- Bayeh, E. (2016), 'The role of empowering women and achieving gender equality to the sustainable development of ethiopia', *Pacific Science Review B: Humanities and Social Sciences* 2(1), 37–42.
- Bradshaw, S. & Howard, P. N. (2018), 'Challenging truth and trust: A global inventory of organized social media manipulation', *The Computational Propaganda Project* 1, 1–26.
- Breza, E., Stanford, F. C., Alsan, M., Alsan, B., Banerjee, A., Chandrasekhar, A. G., Eichmeyer, S., Glushko, T., Goldsmith-Pinkham, P., Holland, K. et al. (2021), 'Effects of a large-scale social media advertising campaign on holiday travel and covid-19 infections: a cluster randomized controlled trial', *Nature medicine* pp. 1–7.
- Brodersen, K. H., Gallusser, F., Koehler, J., Remy, N., Scott, S. L. et al. (2015), 'Inferring causal impact using bayesian structural time-series models', *The Annals of Applied Statistics* 9(1), 247–274.
- Buller, D. B., Pagoto, S., Baker, K., Walkosz, B. J., Hillhouse, J., Henry, K. L., Berteletti, J. & Bibeau, J. (2021), 'Results of a social media campaign to prevent indoor tanning by teens: A randomized controlled trial', *Preventive medicine reports* 22, 101382.
- Christakis, N. A. & Fowler, J. H. (2013), 'Social contagion theory: examining dynamic social networks and human behavior', *Statistics in medicine* 32(4), 556–577.
- Coviello, L., Sohn, Y., Kramer, A. D., Marlow, C., Franceschetti, M., Christakis, N. A. & Fowler, J. H. (2014), 'Detecting emotional contagion in massive social networks', *PloS one* 9(3), e90315.
- Crimson Hexagon (2019). Last accessed 27 December 2019. URL: <https://forsight.crimsonhexagon.com/>

- De Choudhury, M., Jhaver, S., Sugar, B. & Weber, I. (2016), Social media participation in an activist movement for racial equality, in 'Tenth International AAAI Conference on Web and Social Media', AAAI.
- DellaPosta, D., Shi, Y. & Macy, M. (2015), 'Why do liberals drink lattes?', *American Journal of Sociology* 120(5), 1473–1511.
- Dincelli, E., Hong, Y. & DePaula, N. (2016), 'Information diffusion and opinion change during the gezi park protests: Homophily or social influence?', *Proceedings of the Association for Information Science and Technology* 53(1), 1–5.
- Edwards, B. & Marullo, S. (1995), 'Organizational mortality in a declining social movement: The demise of peace movement organizations in the end of the cold war era', *American Sociological Review* pp. 908–927.
- Ferrara, E., Chang, H., Chen, E., Muric, G. & Patel, J. (2020), 'Characterizing social media manipulation in the 2020 us presidential election', *First Monday* 25, 11.
- Fileborn, B., & Loney-Howes, R. (Eds.). (2019). *#MeToo and the politics of social change*. Springer Nature.
- Freischlag, J. A., & Faria, P. (2018). It is time for women (and men) to be brave: a consequence of the #MeToo movement. *Jama* 319(17), 1761-1762.
- Fuks, W. (2020). Pycausalimpact [Source code]. Retrieved from <https://pypi.org/project/pycausalimpact/>
- Gaspar, R., Pedro, C., Panagiotopoulos, P., & Seibt, B. (2016). Beyond positive or negative: Qualitative sentiment analysis of social media reactions to unexpected stressful events. *Computers in Human Behavior*, 56, 179-191.
- Goldberg, A. & Stein, S. K. (2018), 'Beyond social contagion: Associative diffusion and the emergence of cultural variation', *American Sociological Review* 83(5), 897–932.
- He, J., Hong, L., Frias-Martinez, V. & Torrens, P. (2015), Uncovering social media reaction pattern to protest events: a spatiotemporal dynamics perspective of ferguson unrest, in 'International conference on social informatics', Springer, pp. 67–81.
- Hong, L., Yang, W., Resnik, P., & Frias-Martinez, V. (2016, November). Uncovering topic dynamics of social media and news: the case of Ferguson. In *International Conference on Social Informatics* (pp. 240-256). Springer, Cham.
- Hodas, N. O. & Lerman, K. (2014), 'The simple rules of social contagion', *Scientific reports* 4, 4343.
- Lee, S., Ha, T., Lee, D. & Kim, J. H. (2018), 'Understanding the majority opinion formation process in online environments: an exploratory approach to facebook', *Information Processing & Management* 54(6), 1115–1128.
- Leung, R. & Williams, R. (2019), '#metoo and intersectionality: An examination of the #metoo movement through the r. kelly scandal', *Journal of Communication Inquiry* 43(4), 349–371.
- Mønsted, B., Sapiezynski, P., Ferrara, E. & Lehmann, S. (2017), 'Evidence of complex contagion of information in social media: An experiment using twitter bots', *PloS one* 12(9), e0184148.
- Myers, S. A., & Leskovec, J. (2012, December). Clash of the contagions: Cooperation and competition in information diffusion. In *2012 IEEE 12th international conference on data mining* (pp. 539-548). IEEE.
- Pennington, J., Socher, R. & Manning, C. D. (2014), Glove: Global vectors for word representation, in 'Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)', Association for Computational Linguistics, pp. 1532–1543.
- Samuel, J., Rahman, M. M., Ali, G. M. N., Samuel, Y., Pelaez, A., Chong, P. H. J., & Yakubov, M. (2020). Feeling Positive About Reopening? New Normal Scenarios From COVID-19 US Reopen Sentiment Analytics. *IEEE Access* 8, 142173-142190.
- Shao, C., Ciampaglia, G. L., Varol, Ö., Yang, K. C., Flammini, A. & Menczer, F. (2018). The spread of low-credibility content by social bots. *Nature communications* 9(1), 1–9.
- Smith, S. T., Kao, E. K., Shah, D. C., Simek, O., & Rubin, D. B. (2018, June). Influence estimation on social media networks using causal inference. In *2018 IEEE Statistical Signal Processing Workshop (SSP)* (pp. 328-332). IEEE.
- Thompson, A., Hollis, S., Herman, K. C., Reinke, W. M., Hawley, K., & Magee, S. (2020). Evaluation of a social media campaign on youth mental health stigma and help-seeking. *School psychology review* 50(1), 36-41.
- Yoo, E., Gu, B., & Rabinovich, E. (2019). Diffusion on social media platforms: A point process model for interaction among similar content. *Journal of Management Information Systems* 36(4), 1105-1141.

- Yu, X., Daida, S. R., Boy, J., & Hong, L. (2020, October). The Effect of Structural Affinity on the Diffusion of a Transnational Online Movement: The Case of# MeToo. In *International Conference on Social Informatics* (pp. 447-460). Springer, Cham.
- Yu, X., Daida, S. R., Bentula, L., & Hong, L. (2020). Characteristics of information spreading across nations. *Proceedings of the Association for Information Science and Technology* 57(1), e309.

Manika Saha¹, Tom Bartindale¹, Sharifa Sultana², Gillian Oliver¹, Shakuntala Haraksingh Thilsted³, Syed Ishtiaque Ahmed⁴, Delvin Varghese¹, Patrick Olivier¹ (2022): Pictorial Consent: Fieldwork Reflections. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_ep11

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_ep11

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

Pictorial Consent: Fieldwork Reflections

Manika Saha¹, Tom Bartindale¹, Sharifa Sultana², Gillian Oliver¹, Shakuntala Haraksingh Thilsted³, Syed Ishtiaque Ahmed⁴, Delvin Varghese¹, Patrick Olivier¹

¹*Faculty of Information Technology, Monash University, Australia;* ²*Information Science, Cornell University, USA;* ³*WorldFish, Malaysia;* ⁴*Computer Science, University of Toronto, Canada*

Contact author: manika.saha@monash.edu

Abstract. Obtaining participants' informed consent is a fundamental ethical requirement of human-centered research. Researchers working with disadvantaged populations in the Global South face the challenge of communicating to participants the many aspects of a study that require consent, for example, the goals of the research and the data privacy risks and assurances. We reflect on our experience of conducting fieldwork in rural Bangladesh and the inadequacy of standard practices of obtaining written or verbal consent. Consent practices that are deemed effective and sufficient in the Global North have left many of our Bangladeshi participants confused (at best) and indifferent (at worst), thereby jeopardizing the ethical integrity of our research. In response, we developed a pictorial consent communication process. Our reflections on its use have led us to highlight the detachment of traditional consent processes from the realities of the field, and call for their root-and-branch re-evaluation. This is a preliminary contribution, intended to provoke discussion and action on a more inclusive informed consent process design.

1 Introduction

Obtaining the consent of participants is a well-established ethical requirement in any human-centered research study, and this consent must be “informed”, that is, participants can only give consent if they have a clear understanding of what the research is, and what participation in the research involves, including objectives of the research, the activities entailed, and the potential benefits and risks of participation (Kasturiaratch et al., 1999; Moodley and Myer, 2007; Rose et al., 2009). Here "risk" includes not only the immediate physical and psychological consequences of participation, but the risks pertaining to the participants' data being knowingly or unknowingly highlighted, shared or disseminated to people beyond the research team.

Where the capabilities and lived experiences of participants in a given study are far removed from those of researchers, for example, the women of marginalized rural communities with whom we work, achieving informed consent is no small matter. Such women typically have limited literacy, little awareness or interest in the norms and practices of academic research communities, and no understanding of the cloud-based technologies and infrastructures that researchers claim will preserve the integrity and privacy of the data they contribute. Perhaps most alarmingly, many studies have demonstrated how disadvantaged communities in the Global South frequently participate in research projects due to their unconditional faith in outsider researchers, without adequately understanding the research objectives and outcomes (Nishimura et al., 2013; Synnot et al., 2014).

As part of a broader human-centered design research study we conducted a participatory media production activity with marginalized female farmers in rural Bangladesh. Rather than using the traditional way of seeking participants' consent, which involves verbal and written presentations (Sterling and Rangaswamy, 2010), we used hand-drawn pictures to supplement our verbal explanation of key terms. As we explain in the sections that follow, the overall experience was broadly positive, and based on this we propose more widespread adoption of such an approach. Indeed, we contend that researchers in human-computer interaction for development (HCI4D) and information and communication technology for development (ICT4D) are obliged to reexamine the morals and ethics of the methods by which they obtain “informed consent” from participants with low-literacy, particularly where they are under-served, under-resourced and experiencing substantial and systematic deprivation.

While any researcher should be commended for engaging wholeheartedly with the ethical review process, the requirements and expectations of what are historically Western bureaucratic mechanisms are far removed from the realities of field work with many communities in the Global South. This paper documents our reflections on our design of a pictorial informed consent form, and its use in a case study with marginalized female farmers in rural Bangladesh. As a result we describe some ethical dilemmas and questions, and make suggestions, for CSCW,

ICT and HCI "for development" researchers who are seeking to engage their participants on solid ethical grounds.

2 Informed Consent in ICT for Development

Values such as privacy and trust underpin the concept of informed consent, and these require the provision of clear information about the goals of a study, potential benefits and risks of participation, and the mechanisms by which the protection of confidential information provided by a study's participants will be guaranteed. For example, Friedman and colleagues (Friedman et al., 2013, 2000) articulated a conceptual model of informed consent (for online interaction) that had five components: i) entitlement to disclosure of the provision of correct and adequate information about a study; ii) comprehension (i.e., the participant's correct understanding of the given information); iii) voluntary participation without control and coercion; iv) competence (i.e., the participant's possession of the mental, emotional, and physical capabilities necessary to provide informed consent); and v) agreement (i.e., the participant's freedom to accept or decline to participate). Notably, the process includes both comprehension and discourse (Faden and Beauchamp, 1986), and there is a subtle connection between informed consent and knowledge production:

“The information itself on which decisions about research participation are made, along with the concepts through which we imagine its processing, are socially constructed and are about particular forms of knowledge and approaches to understanding or learning” (David et al., 2001) (p.350).

2.1 ICTD, Consent and Marginalised Participants

HCI4D and ICT4D have traditionally focused on undertaking research with marginalized communities, for example, in exploring or designing for health (Kumar and Anderson, 2015; Pai et al., 2013), education (Vashistha and Thies, 2012), agriculture (Patel et al., 2010; Gandhi et al., 2007) and gender relations (Sambasivan et al., 2019). In such studies, understanding local contexts, including social, economic, and political factors, has been at the forefront of researchers' attempts to sustainably use digital technologies in development (Walsham, 2017; Harris, 2016; Avgerou, 2010; Heeks, 2003). However, such engagements are usually markedly different from those with other groups, not just because of the socioeconomic and cultural divides between researchers and participants, but because members of such communities are far less likely to be able to read and write.

Despite this, even in a multidisciplinary field such as ICT4D, just how socio-cultural factors shape informed consent is still under-explored. Pink documented that informed consent is shaped by many social and cultural factors,

including power dynamics, and gatekeepers' roles (Pink, 2012). Indeed, previous research has shown that there is a tendency among participants to want to please researchers (Al-Ameen et al., 2020; Dell et al., 2012). For instance, Dell and colleagues showed in their research that, despite understanding the research objectives and the risks and benefits of participation, participants tended to answer questions according to what they thought the researchers wanted to hear (Dell et al., 2012). Likewise, in previous studies of privacy in Bangladesh, participants did not hesitate to give their personal information to researchers, despite knowing the risks of information leakage (Al-Ameen et al., 2020).

As CSCW, ICT and HCI "for development" researchers, the question of how to obtain informed consent in studies involving marginalized communities is not only an ever-present practical concern (at least for any fieldworker in rural Bangladesh), but it is also an ethical and a moral question that must be addressed. The gap between the processes that we are taught to be sufficient – at least enough for an ethics review board at our Australian university – and the realities on the ground, require us to seek alternative approaches (Flicker et al., 2007; George et al., 2014).

2.2 Visual Communication & Informed Consent

Art-based research studies in social science have established the potential of visual presentations to promote autonomy, and enhance understanding and awareness of research among people with low levels of reading and writing skills (Chilton and Leavy, 2014). In the context of media and communication, visual communication for development and social change was initiated during the 1980s and '90s, and continued in concert with growing technological development, for example, in participatory video making and mobile learning (Prasad, 2020). Moreover, research methods themselves more directly leveraged visual research methods (VRM) (Rose et al., 2009) to leverage participant's critical reflection (Pink, 2012) in an attempt to capture "authentic" community views. One approach in particular, Photovoice (Wang, 1999), is a participatory action research method that continues to be widely deployed to engage marginalized communities to identify, represent and share their strengths, challenges and priorities.

Instructional graphics have long been used within medical science consent process, both to help patients consenting to treatment decisions (Davis et al., 1998) as well as when taking part in health research where simplified forms using instructional graphics are used to explain the treatments and tests to patients. Likewise, the use of audio and visual aids are recognized as valuable tools for engaging and communicating with children across a wide range of disabilities (Stalker and Connors, 2003; Paramasivam et al., 2021). Visual learning aids, including photographs, pictures and videos, are also crucial communication resources in engaging, informing and assisting the learning processes of people with communication and intellectual disabilities (Sperotto, 2016).

In relation to informed consent in development contexts, Lie and Witteveen (Lie and Witteveen, 2017) described their audio-visual informed consent process

in a project based in KwaZulu-Natal in which they used films instead of written forms. While they claimed their approach has a significant impact on participants' comprehension of research and ethical considerations, they also documented several significant issues, including the time-consuming nature of the production process (including translation), and field-related technical barriers such as maintaining the display equipment's battery power. In our case study, we explored an alternative approach that uses pictorial elements to augment the consent process and promote both comprehension and dialogue with the women farmers that were our research participants.

3 Case Study

3.1 Context

The wider research endeavour, within which our reflections on pictorial consent is situated, was a participatory media production activity with marginalized female farmers in rural Bangladesh, many of whom had low or very low levels of literacy. The goal here was for the women to use a prototype technology to capture their daily lives and related needs, challenges, and priorities; and to achieve this with minimal support from the research team. As such, the overall research goal was to study the barriers and limitations of the community production process rather than the visual consent process itself. Our insights correspond to "reflections" on the consent process of the overarching study (although all interactions with participants were recorded in a note and/or audio form), rather than formal "data", and constitute the outcome of researcher reflexivity, rather than what would conventionally be considered the "findings" of a research study that targets the consent process itself.

The overarching study involved 26 rural women farmers who engaged in participatory media production activities in Parbatipur Upazila, Rangpur, Bangladesh. Among the 26 participants, four women were illiterate; 12 women had not finished primary school, and had completed their primary school education. The remaining eight were adolescent upper high school girls who engaged in agricultural activities with their families. In order for us to include everyone in the research study the participants needed to understand all the steps and activities involved in the production of the media, the overall research goals, the data we were collecting and the risks and benefits of participation.

In our ethics application to Monash University Human Research Ethics Committee (MUHREC) we included a traditional two-page explanatory statement and a one-page consent form (translated into Bengali). However, during the field study, and during prior preparatory work in Bangladesh, several questions were raised as to how we would successfully communicate all the necessary issues just using our text-based explanatory statement and consent form (even when read out). In response to these concerns we developed a modified statement and consent form in a graphical format with simplified explanations in the local language (Bengali).

3.2 Designing the Pictorial Consent Form

In the design of our pictorial consent form (see Figure 1) we have used the criteria of: (i) disclosure, (ii) comprehension, (iii) voluntary participation; (iv) competence; and (v) agreement to participate in the research (Friedman et al., 2000). We prepared our consent form to be as concise and direct (i.e., using figurative imagery) as possible to aid comprehension, and avoided jargon by using plain language (Bangla translation from plain English). The first author of the paper (a Bangladeshi woman and native Bengali speaker) sketched pictures of different aspects of the research process and conducted the community research engagement activities including the consent step.

3.3 Doing Pictorial Consent

We used the pictorial consent form as a tool to describe our research aims, the participants' roles, and all ethical considerations related to participant engagement in our study. Our reflections are based on our informal observations of the participants' interactions with the pictorial consent form, what they understood, and what they had difficulty understanding. The women were addressed by the researcher as a group and provided with individual copies of the pictorial consent form. The form was used to introduce the team (i.e., the researchers) and describe our research objectives, the video production activities, the process of anonymous data sharing and storage, withdrawal, the complaints process, and the nature of our relationship and communication with local contacts (through which recruitment had been conducted).

We went through all the criteria of informed consent mentioned above in Section 3.2, using the pictures to aid our explanations and observing the participants' engagement, interactions, and behaviors during the consent-taking procedure. The engagement was friendly and result in spontaneous conversations that went far beyond our previous experiences of the traditional consent activities. We spent around 30 minutes to explain and engage our participants with this pictorial consent forms. At the end of this time, participants choose verbal or signed consents according to their preference. The verbal consents were recorded via a tape-recorder device, and we collected signed consent forms after the informed consent taking process. All observations, informal discussion, and note-taking were performed upon receiving permission from the participants. Our experience of developing the pictorial consent form, observations, and field notes helped to articulate our research findings and have been documented in this paper as reflections. It is worth noting that this pictorial consent form was not intended to be used "stand alone", i.e., without any verbal explanation or simplified written description. Hence, we did not expect our pictures to wholly represent what in many cases were complex concepts of data, participation and risk etc. Instead we used the visual consent form to scaffold facilitation of dialogue with participants.

Excerpt from visual consent form		Information (English translation)	
Visual	আমার নাম এবং আমি	Information (English translation)	আমার নাম এবং আমি
	বাগদেশের মহিলাদের মধ্যে কৃষি, স্বাস্থ্য ও পুষ্টি এবং ডিজিটাল প্রযুক্তির ব্যবহার সম্পর্কে কিছু বিষয় বোঝার ও অনুসন্ধান করতে চাইব। এই শিটটি আপনাকে আমাদের কাজের লক্ষ্য সম্পর্কে জানতে পাবে এবং আপনার সাথে আমাদের এই কর্মকলাপ সম্পর্কে জানাবে।	My name is _____ and I am a _____ here to understand and investigate a few things on agriculture, health and nutrition and uses of digital technology among the women in a rural context in Bangladesh. This sheet will tell you about our goal of the industry and engagement activities with you.	আমাদের সাথে আমাদের জড়িত হওয়ার উদ্দেশ্য কি? - আমরা আপনার নিয়ে ভিডিও তৈরি করব যেখানে আপনার কৃষি, স্বাস্থ্য, পুষ্টি এবং প্রযুক্তির ব্যবহার সম্পর্কে আপনার মতামত গুরুত্বপূর্ণ হবে এবং আমরা সেগুলোর ভিডিও বানাতে। - আমরা আপনার ভূমিকা এবং আমাদের কী কী করতে হবে তা বলব।
	আপনার কাছ থেকে আমাদের এই সকল তথ্য জানার উদ্দেশ্য কি? - কৃষি সংক্রান্ত আপনার অভিজ্ঞতা এবং এই ক্ষেত্রে আপনার যেসব সমস্যা পড়তে হয় তা বোঝার জন্য।	What is our aim of understanding from you? - To understand your experience, constraints regarding agriculture.	আমাদের সাথে আমাদের জড়িত হওয়ার উদ্দেশ্য কি? - আমরা আপনার সাথে নিয়ে কিছু গুরুত্বপূর্ণ ও উঠান বৈঠক করব।
	আপনার অবস্থান বোঝার জন্য, আপনার খাদ্যাভ্যাস, স্বাস্থ্য ও পুষ্টি এবং এইসব ক্ষেত্রের আশেপাশে ডিজিটাল প্রযুক্তি (ICT) ব্যবহার কেমন তা জানব।	To understand your situation, understanding and practice regarding your food habits, health and nutrition and using digital technology/ Information and Communication Technology (ICT) around these areas.	- এই গুরুত্বপূর্ণ গুলোতে আপনার নিয়ে কিছু কার্যকলাপ থাকবে। যেখানে বিভিন্ন বিষয় যেমন-কৃষি, স্বাস্থ্য, পুষ্টি এবং প্রযুক্তির অভিজ্ঞতা নিয়ে আলোচনা করা হবে। - আপনার দৈনন্দিন জীবনের কৃষি, স্বাস্থ্য এবং পুষ্টি নিয়ে গুরুত্বপূর্ণ গল্পগুলো সম্পর্কে আমরা জানতে। - আমরা একটি স্টোরী বোর্ড বানাতে যাতে আপনি আপনার মতামত প্রকাশ করতে পারেন। - আমরা একটি স্টোরী বোর্ড বানাতে যাতে আপনি আপনার মতামত প্রকাশ করতে পারেন। - আমরা একটি স্টোরী বোর্ড বানাতে যাতে আপনি আপনার মতামত প্রকাশ করতে পারেন।
	একজন মহিলা হিসেবে পরিবার এবং সমাজে আপনার ভূমিকা সম্পর্কে জানা।	To understand your role as a woman in your family and community.	আমাদের সাথে আমাদের জড়িত হওয়ার উদ্দেশ্য কি? - আমরা আপনার সাথে নিয়ে কিছু গুরুত্বপূর্ণ ও উঠান বৈঠক করব।
	আমরা কিছু অডিও রেকর্ড করতে চাই। আপনি আমাদের সাথে আলোচনা করে আমাদের মতামত জানতে পারবেন।	We would like to audio record our discussions so we can listen to them later again when we go back from the field. We will ask for your consent for making the participatory videos with you.	আমাদের সাথে আমাদের জড়িত হওয়ার উদ্দেশ্য কি? - আমরা আপনার সাথে নিয়ে কিছু গুরুত্বপূর্ণ ও উঠান বৈঠক করব।
	গোপনীয়তা এবং ফলাফল/তথ্য শেয়ার আমরা এই অনুসন্ধানের একটি রিপোর্ট তৈরি করতে চাই। আপনি আমাদের সাথে আলোচনা করে আমাদের মতামত জানতে পারবেন।	Confidentiality and share the findings/ information with others Part of our investigation is to tell other people what we have done. We will write a report about our understanding and findings. In this report we may use quotes of what you said. If we do this, your name and your details will be removed so no one will know about you. Your name or any identifying information will not be shared with anyone and anywhere.	আমাদের সাথে আমাদের জড়িত হওয়ার উদ্দেশ্য কি? - আমরা আপনার সাথে নিয়ে কিছু গুরুত্বপূর্ণ ও উঠান বৈঠক করব।
	এই রিপোর্ট, অডিও এবং ভিডিও গুলো পরবর্তীতে বিভিন্ন সামাজিক ফোরামে আমরা মেম্বার হিসেবে উপস্থিত হব এবং বিভিন্ন পরেখা পরেখা প্রক্রিয়ায় অডিও এবং ভিডিও গুলো উপস্থিত হব। আমরা আপনার গোপনীয়তা এবং সুরক্ষা নিশ্চিত করব।	Findings and outcomes of the research will be published in journal articles and conferences. We may use transcriptions, audio and video from the recordings to develop stories and reports, but none of these will contain your names or any identifying data. The photos and videos may be used in the reports to the stakeholders and could share in the social media such as in Facebook and YouTube but there will be no information about you.	আমাদের সাথে আমাদের জড়িত হওয়ার উদ্দেশ্য কি? - আমরা আপনার সাথে নিয়ে কিছু গুরুত্বপূর্ণ ও উঠান বৈঠক করব।
	সম্মতি যদি আপনি অংশগ্রহণ করতে ইচ্ছুক হন তবে আমাদের আপনার একটি অনুমতি দিয়ে স্বাক্ষর করুন। আপনি যদি স্বাক্ষর করতে না চান তবে আমরা আপনার মুখে রেকর্ড করে নিব।	Consent If you want to take part we will ask to sign a form giving your permission. You could give us your verbal consent as well if you prefer instead of signing consent form. We will record your verbal consent in the tape-recorder.	আমাদের সাথে আমাদের জড়িত হওয়ার উদ্দেশ্য কি? - আমরা আপনার সাথে নিয়ে কিছু গুরুত্বপূর্ণ ও উঠান বৈঠক করব।
	প্রশ্ন এবং উত্তর আপনার কোন প্রশ্ন থাকলে চাইলে প্রশ্ন উত্তর সেশনে আমাদের সাথে আলোচনা করে জানতে পারবেন।	Questions and answers You can ask any questions before then or during the allocated Q&A session.	আমাদের সাথে আমাদের জড়িত হওয়ার উদ্দেশ্য কি? - আমরা আপনার সাথে নিয়ে কিছু গুরুত্বপূর্ণ ও উঠান বৈঠক করব।
	অংশগ্রহণ থেকে প্রত্যাহার আপনার মন পরিবর্তন করতে পারেন। আপনি আমাদের সাথে আলোচনা করে আমাদের মতামত জানতে পারবেন।	Withdrawing from the participation You can change your mind about taking part. You will only be requested to respond to discussion questions relevant to your interest. If you face any kind of discomfort, you can leave the discussion, training, workshop and video activities at any time. You will be free to withdraw from this participation at any stage and you do not have to give a reason. Participation in these activities are voluntary and you are under no obligation to consent to participation.	আমাদের সাথে আমাদের জড়িত হওয়ার উদ্দেশ্য কি? - আমরা আপনার সাথে নিয়ে কিছু গুরুত্বপূর্ণ ও উঠান বৈঠক করব।
	আপনার যদি কোনো প্রশ্ন থাকে যা আপনি আমাদের সাথে আলোচনা করতে চান, আপনি আমাদের সাথে আলোচনা করে আমাদের মতামত জানতে পারবেন।	If you have any questions that you wish to discuss, you can communicate with us. You can talk to one of our representatives from _____ in your community level and then, we can communicate with you directly. Your community level _____	আমাদের সাথে আমাদের জড়িত হওয়ার উদ্দেশ্য কি? - আমরা আপনার সাথে নিয়ে কিছু গুরুত্বপূর্ণ ও উঠান বৈঠক করব।
	এই অনুসন্ধান পরিকল্পনাটি 'রিসার্চ ইথিক্স কমিটি' দ্বারা পর্যালোচনা করা হয়েছে এবং এই পরিকল্পনাটি সুরক্ষিত এবং এই পরিকল্পনাটি সুরক্ষিত এবং এই পরিকল্পনাটি সুরক্ষিত।	This investigation plan has been checked by the Research Ethics Committee which ensure all that all activities and this research project are fair and safe. This research project has been checked by the Research Ethics Committee in _____	আমাদের সাথে আমাদের জড়িত হওয়ার উদ্দেশ্য কি? - আমরা আপনার সাথে নিয়ে কিছু গুরুত্বপূর্ণ ও উঠান বৈঠক করব।
	যদি আমাদের এই কার্যক্রম সম্পর্কে আপনার কোন অভিযোগ থাকে, তবে আপনি আমাদের সাথে আলোচনা করে আমাদের মতামত জানতে পারবেন।	If you have any complaints about any aspect of our engagement with you, please contact with _____ representative in your community level. Your community level _____ They can help you to reach other persons in _____ as well as _____	আমাদের সাথে আমাদের জড়িত হওয়ার উদ্দেশ্য কি? - আমরা আপনার সাথে নিয়ে কিছু গুরুত্বপূর্ণ ও উঠান বৈঠক করব।

Figure 1. Examples from the pictorial consent form in which information about the researchers, objectives of the research, field activities, and use of digital technologies for capturing videos were explained using pictures to aid comprehension. The third column is an English translation of the simplified Bangla descriptions included on the pictorial consent form.

4 Reflections from Field Observations and Discussion

4.1 Designing Consenting Mechanisms for Meaningful Interactions

We observed several benefits to using our pictorial consent forms with participants. When introducing the research team and the project to participants, we explained the purpose of the research and our intention to use the visual information and consent form to help them decide whether they wished to proceed with participation in the research. We then provided, using a combination of our visual consent form and verbal instructions, a detailed explanation of project activities related to community production of participatory media. At the conclusion of the project, we asked for participant feedback on the pictorial information and consent sheet. Our participants commented that the visual nature of the consent form was helpful in understanding the motivation and methods of the project. To ensure that this was not just a form of participant response bias, we asked questions of our participants about the research team's intentions and future actions, to which the participants responded: *"Yes, we understood why you came here"*, and they described their engagement activities with us, *"we need to share information regarding our daily life activities that relates to our agriculture, health and nutrition, experience as women and using technologies"* (from field notes during observation).

Furthermore, we discovered to our surprise that the participants found the process of giving informed consent enjoyable. Our participants were very curious and initiated a number of discussions with our team and their fellow participants. For instance, we found them stopping to think out loud about how activities had been described, and that they paid focused attention to certain pictures, turning back and forth to other pictures. Female participants in particular were highly motivated to explore how they were represented. The women highlighted that *"it is interesting to see the pictures specially our pictures"* (from field notes during observation). During our discussions, the women pointed out that they felt that they could relate to the drawings of women as being of rural women, a label with which they strongly identified. They interrogated us about different aspects of the pictures to better understand their involvement in activities related to participatory media production. After we explained some of the activities involved, the facilitator asked what they thought about this pictorial sheet and a typical response was, *"yes, it is easy to understand that we will do some activities to share our experiences and then we will take videos"* (from field notes during observation). Moreover, as was our intention, we noted that our pictorial consent form was particularly helpful for engaging low-literate people from marginalized communities and discussing non-trivial aspects of research activities, such as participatory media production with digital technologies.

In our field observations, we recognized that our pictorial consent form helped us describe our research ideas, objectives, participant engagement activities, and some ethical considerations (such as consent for recording their conversations, withdrawing from research, and further communication with us and local

authorities). Indeed, we found this experience more straightforward than our previous experiences (without the use of visual consent forms). We found that visual consent also helped us to refine our communications with the community before commencing the research related activities. Many practitioners we have worked with in past have noted that "*community entry*" is a fraught process that needs to be handled sensitively. Hence, we posit that an alternative consent taking process such as pictorial consent could be used as a tool for an introductory engagement activity with participants to rapport and make better interaction for research and fieldwork with low-literate communities.

4.2 Ethical Dilemmas and Challenges

So far we have described how our pictorial consent form assisted us in explaining our project to the communities involved. We have also discussed some of the inherent ethical considerations. However, by slowing down the consenting process and spending meaningful time with communities through the activities, we discovered that communities had limited understanding on *why* consent was required in the first place. Why was it, they wondered, necessary to obtain consent and explain benefits and risks of a project before participating in it in the first place? Perhaps this was due to their previous interactions with institutional stakeholders that they were not prepared for the process of consent. Regardless of the underlying reasons for this reaction, we faced challenges in communicating simply about the ethical considerations and decisions that participants need to make before committing to proceed with this research. To be sure, despite the voluntary nature of this (and any research), the decision to initiate participation is considered an important decision. For instance, one of the risks of participating in any research is the security and privacy of their identity information and the wider dissemination of participants' information. Though we tried to explain these ethical issues in simple language and to use our visual aids, our participants did not clearly understand some ethical concepts such as where and how (e.g., academic journals and reports) their information would be used. We recognise that it is challenging for individuals without prior experience to conceptualise what journal papers, reports and conferences are, how their information will be presented within those venues, and the direct benefits, harms and risks for them related to these processes. They also had little understanding of the value of knowing where we would store their information. For instance, what did it mean to store information in a computer or a hard disc, and the notion that it was securely stored was another added complexity to this fact. Though we verbally explained this in simple language (as best as we could) and used pictures for clarification, we found it was difficult for them to understand such ethical considerations due to their limited education and lack of prior exposure. A key consideration here is that the problem lies not just in how it is presented (i.e. in text or picture formats), but in the inherent complexity of traditional ethical requirements. Indeed, it could be argued that a special literacy exists around ethical conduct of research or practice that

requires specialised training to effectively navigate, both as an administrator and participant of ethics practices.

These dilemmas and challenges on taking consent via traditional ethics mechanisms raised several questions. We also wanted to understand motivations for participation in our research when invited, as cultural expectations dictate that guests' expectations should be met, i.e. researchers asking for consent should be affirmed. Furthermore, in a patriarchal society like Bangladesh, women are marginalized and deprived, as most of their fundamental rights are unmet (Quisumbing, 2012; Kabeer, 2016, 1997). Women's suppression is ensconced within all aspects of women's lives in Bangladesh, including in their family, community, and society, and it impacts their social, health, and economic status (Kabeer, 1997). Women, especially in rural Bangladesh, have limited freedom to argue, protest, complain, share their desires, and make decisions in their lives (Sultana, 2010). Moreover, decisions about daily activities (e.g., household chores and going outside) and crucial life-changing matters (e.g., starting or stopping education, getting married, and having children) are typically made by women's families without the consent of the woman herself (Sultana et al., 2018). Hence, the women in our study were culturally unaccustomed to saying "no" and had limited agency to make their own decisions. They value other people's wishes more than their own.

We understand that in a conservative and patriarchal society that does not ask these marginalized women's consent regarding crucial life events, requesting their consent to take part in a research study raises the following question: How much consent is relatable according to their actual lived experience? Where consent is not obtained from women regarding their important life matters, it might be absurd for the participants to ask permission to find out information regarding their participation in a research study. Moreover, previous research has shown that there is a common tendency in postcolonial cultures to please outsiders (Ahmed et al., 2017). Specifically, there is a common tendency to say "yes" to requests for participation and, generally, to provide positive answers to researchers (Dell et al., 2012). One of the informal comments from our discussion with the participants reinforced this tendency as a common practice among disadvantaged communities. *"We give our consent to whoever comes to us and asks to take part in their work. It is not good to say 'No' to the outsider who comes to us from a long distance" (from field notes during observation).* We noticed that the participants felt some obligation to please the researcher, which was more important to them than understanding the risks and benefits of participating in the research. The participants willingly participated in our media production activities.

Furthermore, we noticed that participants attributed more importance to money and other material rewards for participation than their consent to participate in the study. One of the motives for these marginalized women to say "yes" and give their consent was compensation from researchers for their participation in such forms as money, food, or gifts. The participants were from a disadvantaged group that struggles daily to meet basic requirements, such as food, due to a lack of

money. Hence, receiving money or food by participating in research activities is more important than assessing the risks and benefits of participating in research for broader social goods. For instance, we verbally asked them why they had consented to participate in our research. Several participants commented that they took part because they were motivated by compensation rather than the purpose of the research. One of the informal comments from our participant was, "*we do not see any harm to take part in while sometimes there are benefits as they provide us food or money*" (from field notes during observation). Hence, our observation reflects conflicting priorities between researchers and marginalized participants, as the general goals of researchers are knowledge production and contribution to broader development. In contrast, a woman from a marginalized community prioritizes meeting her basic needs.

4.3 A Manifesto for Improving the Informed Consent Process

Obtaining informed consent from participants is the cornerstone of research conducted within academia. However, this practice originated from Western-centered bureaucratic mechanisms and is often overseen by Institutional Review Boards (IRB) or ethics boards that scrutinize human studies to minimize risks and ensure the quality of research data (Cseko and Tremaine, 2013). There is an implicit assumption that translation or verbal communication of this material would be equivalent to the written form and therefore be sufficient. Our fieldwork foregrounded ethical questions to all development researchers about how disadvantaged communities understand research regarding consent, privacy, the security of their data, information dissemination, and other risks and benefits of participating in a research study. We strongly argue that the current traditional informed consent taking process with limited literacy people is broken. We commend the research community for engaging wholeheartedly with the ethical review process. However, we posit that the requirements and expectations of these Western centred bureaucratic mechanisms are significantly detached from the realities of fieldwork with many communities.

Hence, we raised concerns about designing informed consent for marginalized people. We tried to design an "informed consent" process. However, then we realized that we would also need a voiced consent whereby participants would fully share their views and opinions by understanding the benefits and risks of participating in a research study. We need their full participation to ensure rigour in the research. Hence, we are arguing for "informed and voiced consent," as we believe "informed consent" is not enough to inform them about a study's benefits and risks; it is also essential to understand their actual perspectives. Based on our experience, they may be willing to take the risk of participating in research. After all, they value basic needs (e.g., money and food) and want gifts. Therefore, researchers must adequately consider marginalized participants' values when conducting research; for example, we can consider participants' priority incentives or foods that need to be provided in similarly marginalized communities.

Moreover, some dilemmas were around how to explain complex ethical consideration concepts such as data storage, privacy, usage of participants' information, and ethical boards. Hence, we suggest that researchers take full responsibility for taking care of all these ethical considerations with a robust moral intention. We think it should be essential to researchers' ethical and moral considerations to refrain from harming marginalized communities, in which people typically have a limited understanding of complex research. Therefore, we think there are plenty of potential areas for the researcher demographic that need to explore conducting informed consent with marginalized communities. We also argue that participants' socio-cultural contexts should consider designing an informed consent-taking process before conducting at the field level with disadvantaged people.

We want to question how technologies can help marginalized people to provide "informed and voiced" consent to allow for rigorous studies that advance community development. In these challenging situations, technology can play a role in finding alternative approaches (such as using media) for engaging and conducting the informed consent process in a more appropriate and meaningful way where the above-raised issues would be addressed. A re-evaluation of the Western understanding of ethics, centred around IRBs and ethics committees, is needed to bridge the gaps in the realities of fieldwork in these contexts. We recommend that ICT and HCI for development researchers come forward, do more research, and explore better designs and approaches to establish an effective informed consent methodology. We are also calling for more research, specifically action-based research, to explore and find solutions to informed to voiced consent design for marginalized populations. We advocate for informed and voiced consent design with communities to find better ways of obtaining informed consent. We also encourage to explore pictorial visual consent forms and other alternative approaches to take consent from research participants where both researchers and participants' interests will meet with clear mutual understanding. Thus, we can ensure we keep our moral obligations to do no harm.

5 Conclusion and Future Work

We reflect that simplifying informed consent forms by adding pictures makes these forms more engaging for participants. Moreover, these pictures made it easier for us to explain the activities involved in our participatory media production research study. However, our study indicates that there is still a lack of comprehension of different aspects of the associated ethical considerations, such as the benefits and risks of participation in a study. Indeed, our study shed further light on how these disadvantaged participants' had been shaped by participants' values and socio-cultural norms . Therefore, we advocate for more research and design work with such participants and the development of an effective process to increase comprehension of ethical issues in research studies.

We also argue that, traditional the informed consent process in such contexts is broken, and call for a re-evaluation of ethical review processes to bridge the gaps between the realities in the field and Western concepts of the IRB and acceptable practice. We emphasize the need for an overarching conceptual and methodological approach with an in-depth understanding of informed consent for disadvantaged groups to handle the actual value of ethics for design work. Based on our experiences, we suggest exploring more and developing situated and sustainable strategies for marginalized people to make informed consent decisions when taking part in research in Bangladesh and other developing countries. We encourage CSCW, ICT and HCI for development research communities to expand inclusive informed consent-related work, including design work in different domains and the exploration of alternative approaches to negotiating and obtaining consent from disadvantaged communities.

References

- Ahmed, S. I., M. R. Haque, J. Chen, and N. Dell (2017): 'Digital privacy challenges with shared mobile phone use in Bangladesh'. *Proceedings of the ACM on Human-Computer Interaction*, vol. 1, no. CSCW, pp. 1–20.
- Al-Ameen, M. N., T. Tamanna, S. Nandy, M. M. Ahsan, P. Chandra, and S. I. Ahmed (2020): 'We Don't Give a Second Thought Before Providing Our Information: Understanding Users' Perceptions of Information Collection by Apps in Urban Bangladesh'. In: *Proceedings of the 3rd ACM SIGCAS Conference on Computing and Sustainable Societies*. pp. 32–43.
- Alaei, M., A. Pourshams, N. Altaha, G. Gogiani, and E. Jafari (2013): 'Obtaining informed consent in an illiterate population'. *Middle East journal of digestive diseases*, vol. 5, no. 1, pp. 37.
- Avgerou, C. (2010): 'Discourses on ICT and development'. *Information technologies and international development*, vol. 6, no. 3, pp. 1–18.
- Bashir, M., C. Hayes, A. D. Lambert, and J. P. Kesan (2015): 'Online privacy and informed consent: The dilemma of information asymmetry'. *Proceedings of the Association for Information Science and Technology*, vol. 52, no. 1, pp. 1–10.
- Chilton, G. and P. Leavy (2014): 'Arts-based research practice: Merging social research and the creative arts'. *The Oxford handbook of qualitative research*, pp. 403–422.
- Cseko, G. C. and W. J. Tremaine (2013): 'The role of the institutional review board in the oversight of the ethical aspects of human studies research'. *Nutrition in Clinical Practice*, vol. 28, no. 2, pp. 177–181.
- David, M., R. Edwards, and P. Alldred (2001): 'Children and school-based research: 'informed consent' or 'educated consent'?'. *British educational research journal*, vol. 27, no. 3, pp. 347–365.
- Davis, T. C., H. J. Berkel, R. F. Holcombe, S. Pramanik, and S. G. Divers (1998): 'Informed consent for clinical trials: a comparative study of standard versus simplified forms'. *JNCI: Journal of the National Cancer Institute*, vol. 90, no. 9, pp. 668–674.

- Dell, N., V. Vaidyanathan, I. Medhi, E. Cutrell, and W. Thies (2012): "' Yours is better!' participant response bias in HCI'. In: *Proceedings of the sigchi conference on human factors in computing systems*. pp. 1321–1330.
- Faden, R. R. and T. L. Beauchamp (1986): *A history and theory of informed consent*. Oxford University Press.
- Flicker, S., R. Travers, A. Guta, S. McDonald, and A. Meagher (2007): 'Ethical dilemmas in community-based participatory research: recommendations for institutional review boards'. *Journal of Urban Health*, vol. 84, no. 4, pp. 478–493.
- Friedman, B., E. Felten, and L. I. Millett (2000): 'Informed consent online: A conceptual model and design principles'. *University of Washington Computer Science & Engineering Technical Report 00–12–2*, vol. 8.
- Friedman, B., P. H. Kahn, A. Borning, and A. Hultgren (2013): 'Value sensitive design and information systems'. In: *Early engagement and new technologies: Opening up the laboratory*. Springer, pp. 55–95.
- Gandhi, R., R. Veeraraghavan, K. Toyama, and V. Ramprasad (2007): 'Digital green: Participatory video for agricultural extension'. In: *2007 International conference on information and communication technologies and development*. pp. 1–10.
- George, S., N. Duran, and K. Norris (2014): 'A systematic review of barriers and facilitators to minority research participation among African Americans, Latinos, Asian Americans, and Pacific Islanders'. *American journal of public health*, vol. 104, no. 2, pp. e16–e31.
- Harris, R. W. (2016): 'How ICT4D research fails the poor'. *Information Technology for Development*, vol. 22, no. 1, pp. 177–192.
- Heeks, R. (2003): 'Most eGovernment-for-development projects fail: how can risks be reduced?'. <https://indaba.dev/guide/> (2019): 'Indaba Guide'.
- Kabeer, N. (1997): 'Women, wages and intra-household power relations in urban Bangladesh'. *Development and Change*, vol. 28, no. 2, pp. 261–302.
- Kabeer, N. (2016): 'Gender equality, economic growth, and women's agency: the "endless variety" and "monotonous similarity" of patriarchal constraints'. *Feminist Economics*, vol. 22, no. 1, pp. 295–321.
- Kasturiaratch, N., R. Lie, and J. Seeberg (1999): 'Health ethics in South-East Asia, Vol I: Health ethics in six SEAR countries'. Technical report, WHO Regional Office for South-East Asia.
- Kumar, N. and R. J. Anderson (2015): 'Mobile phones for maternal health in rural India'. In: *Proceedings of the 33rd annual acm conference on human factors in computing systems*. pp. 427–436.
- Lie, R. and L. Witteveen (2017): 'Visual informed consent: Informed consent without forms'. *International Journal of Social Research Methodology*, vol. 20, no. 1, pp. 63–75.
- Minkler, M. (2004): 'Ethical challenges for the "outside" researcher in community-based participatory research'. *Health Education & Behavior*, vol. 31, no. 6, pp. 684–697.
- Mitchell, C. (2011): *Doing visual research*. Sage.

- Moodley, K. and L. Myer (2007): 'Health research ethics committees in South Africa 12 years into democracy'. *BMC Medical Ethics*, vol. 8, no. 1, pp. 1–8.
- Nishimura, A., J. Carey, P. J. Erwin, J. C. Tilburt, M. H. Murad, and J. B. McCormick (2013): 'Improving understanding in the research informed consent process: a systematic review of 54 interventions tested in randomized control trials'. *BMC medical ethics*, vol. 14, no. 1, pp. 1–15.
- Nova, F. F., M. R. Rifat, P. Saha, S. I. Ahmed, and S. Guha (2019): 'Online sexual harassment over anonymous social media in Bangladesh'. In: *Proceedings of the Tenth International Conference on Information and Communication Technologies and Development*. pp. 1–12.
- Pai, N., P. Supe, S. Kore, Y. Nandanwar, A. Hegde, E. Cutrell, and W. Thies (2013): 'Using automated voice calls to improve adherence to iron supplements during pregnancy: a pilot study'. In: *Proceedings of the Sixth International Conference on Information and Communication Technologies and Development: Full Papers-Volume 1*. pp. 153–163.
- Paramasivam, A., A. Jaiswal, R. Minhas, W. Wittich, and R. Spruyt-Rocks (2021): 'Informed Consent or Assent Strategies for Research With Individuals With Deafblindness or Dual Sensory Impairment: A Scoping Review'. *Archives of rehabilitation research and clinical translation*, vol. 3, no. 2, pp. 100115.
- Patel, N., D. Chittamuru, A. Jain, P. Dave, and T. S. Parikh (2010): 'Avaaj otalo: a field study of an interactive voice forum for small farmers in rural india'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 733–742.
- Pink, S. (2012): 'The visual in ethnography: photography, video, cultures and individuals'. *SAGE visual methods*. Thousand Oaks, CA: Sage, pp. 123–144.
- Prasad, K. (2020): 'Communication policy for women's empowerment: Media strategies and insights'. *Handbook of Communication for Development and Social Change*, pp. 519–530.
- Quisumbing, A. (2012): 'Innovative approaches to gender and food security: changing attitudes, changing behaviours'. *IDS Insights*, no. 82.
- Rao, S. M. and B. Gagie (2006): 'Learning through seeing and doing: Visual supports for children with autism'. *Teaching exceptional children*, vol. 38, no. 6, pp. 26–33.
- Rogers, C. G., J. E. Tyson, K. A. Kennedy, R. S. Broyles, and J. F. Hickman (1998): 'Conventional consent with opting in versus simplified consent with opting out: an exploratory trial for studies that do not increase patient risk'. *The Journal of pediatrics*, vol. 132, no. 4, pp. 606–611.
- Rose, G. (2014): 'On the relation between 'visual research methods' and contemporary visual culture'. *The sociological review*, vol. 62, no. 1, pp. 24–46.
- Rose, S., M. Aburto, J. Hagemann, and D. Shahnazarian (2009): 'Informed consent in human subjects research'. *University of South California*. [Online]: Available at: <https://opr.usc.edu/files/2017/04/Informed-Consent-Booklet-4.4>, vol. 13.
- Sambasivan, N., A. Batool, N. Ahmed, T. Matthews, K. Thomas, L. S. Gaytán-Lugo, D. Nemer, E. Bursztein, E. Churchill, and S. Consolvo (2019): '" They Don't Leave Us Alone Anywhere We Go" Gender and Digital Abuse in South Asia'. In: *proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. pp. 1–14.
- Sperotto, L. (2016): 'The visual support for adults with moderate learning and communication disabilities: how visual aids support learning'. *International Journal of Disability, Development and Education*, vol. 63, no. 2, pp. 260–263.

- Stalker, K. and C. Connors (2003): 'Communicating with disabled children'. *Adoption & Fostering*, vol. 27, no. 1, pp. 26–35.
- Sterling, S. R. and N. Rangaswamy (2010): 'Constructing informed consent in ICT4D research'. In: *Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development*. pp. 1–9.
- Sultana, A. (2010): 'Patriarchy and women s subordination: a theoretical analysis'. *Arts Faculty Journal*, pp. 1–18.
- Sultana, S. and S. I. Ahmed (2019): 'Witchcraft and hci: Morality, modernity, and postcolonial computing in rural bangladesh'. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. pp. 1–15.
- Sultana, S., F. Guimbretière, P. Sengers, and N. Dell (2018): 'Design within a patriarchal society: Opportunities and challenges in designing for rural women in bangladesh'. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. pp. 1–13.
- Synnot, A., R. Ryan, M. Prictor, D. Fetherstonhaugh, and B. Parker (2014): 'Audio-visual presentation of information for informed consent for participation in clinical trials'. *Cochrane Database of Systematic Reviews*, no. 5.
- Vashistha, A. and W. Thies (2012): '{IVR} Junction: Building Scalable and Distributed Voice Forums in the Developing World'. In: *6th USENIX/ACM Workshop on Networked Systems for Developing Regions ({NSDR} 12)*.
- Vines, J., R. Clarke, P. Wright, J. McCarthy, and P. Olivier (2013): 'Configuring participation: on how we involve people in design'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 429–438.
- Walsham, G. (2017): 'ICT4D research: reflections on history and future agenda'. *Information Technology for Development*, vol. 23, no. 1, pp. 18–41.
- Wang, C. C. (1999): 'Photovoice: A participatory action research strategy applied to women's health'. *Journal of women's health*, vol. 8, no. 2, pp. 185–192.

Mohallick, I., Monteiro, E, Østerlie, T (2022): The Design Multiple: Sharedness and Multiplicity in Common Information Spaces. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centered Computing on the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_ep12

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_ep12

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

The Design Multiple: Sharedness and Multiplicity in Common Information Space

Itishree Mohallick, Eric Monteiro & Thomas Østerlie

Department of Computer Science, Norwegian University of Science and Technology, Norway

itishree.mohallick@ntnu.no, eric.monteiro@ntnu.no & thomas.osterlie@ntnu.no

Abstract. Common Information Space (CIS) seeks to support communication, interaction, and sharing information to collectively perform work tasks and organize dependencies among the involved actors. The concerted efforts that go into collaborative work require some degree of shared understanding among the actors. However, exactly how much shared understanding is needed remains contested. Our findings from a case study of the complex-product design show that sharedness ranges from significant and moderate to minimal, which is already well established in the CSCW literature. This paper discusses these conventions concerning the CIS to improve collaboration. The case study focuses on the CIS conventions observed in the early-stage design of subsea equipment in an Oil and Gas project. Besides, it confers the significance of such CIS in supporting collaborative work practices. Our research represents an under-researched version of CIS against the dominance of moderate (and significant) degrees of overlap. Relative to the CSCW discourse on CIS, however, the most interesting aspect of our analysis is the presence of the third (“minimal”) overlap. The reported empirical insights can help the researchers to discuss the design-related work practices- sharedness and multiplicity- in a context not well explored in literature.

Introduction

Collaborative work, by definition, is collective. It thus involves communicating, interacting, and sharing information to collectively perform work tasks and organize dependencies among the involved actors in ways the field of CSCW (Computer Supported Cooperative Work) is all about (Schmidt and Bannon, 2013). The joint, concerted efforts that go into collaborative work require a degree of shared understanding among the actors, without which communication thus collaboration would degenerate into a private language-like situation (Wittgenstein 2010) hence cease to function communicatively.

However, exactly *how* much shared understanding is required remains contested with analytic positions ranging from assuming shared understanding to be significant (Munkvold and Ellingsen, 2007; Beretelsen and Brødker, 2001), moderate (Levina and Vast, 2005; Fields et al., 2003), or but minimal (Mol, 2003; Oborn et al., 2011) in the overlap. The theoretical question of degree of shared understanding has strong and immediate implications for empirical and practical collaboration given that studies have reported on challenges of achieving a high degree of shared understanding due to a variety of circumstances, including fragmented, non-integrated systems (Ellingsen and Monteiro, 2003), political differences (Carlile, 2004) and differences in communities of practices that the actors belong to (Hildreth and Kimble 2004). We thus ask the research question: What degree of shared understanding is required for collaborative work?

Our empirical case is particularly well suited to address our research question (Flyvbjerg, 2006). We analyze the collaborative work during the design process of complex equipment within the offshore oil and gas industry in North Europe. This equipment is part of advancing so-called subsea oil and gas production, i.e., equipment part of the 50 meters x 50 meters metal installation residing at the bottom of the sea at depths 200 - 1000 meters with valves, chokes, pipes, measurement sensors, and hydraulic control systems allowing unmanned, remotely operating oil and gas production. On the Norwegian Continental Shelf, already more than 50% of oil and gas production is by such unmanned.

The challenge with designing and manufacturing this equipment hinges on the way “every [oil] well is unique” (National Commission, 2011), i.e., that each designed piece of equipment needs to be tailored to the characteristics of that specific oil well, the details of the geological circumstances as well as access to nearby, existing production facilities. In short, the design is challenging because there is no given template to start from, there are both engineering, material, geological and economic concerns to account for, and the design involves extensive

dialogue between the vendor and the client (an oil operator). We specifically focus on the early-stage design part of the process as this is when design ideas, visions, and prototypes have yet to settle, thus maximizing the likelihood of challenges with shared understanding across the different professional and organizational actors involved.

In addition, we study the notion of multiplicity as a concept while investigating the early-stage design work for a project (Mol, 2003). Multiplicity is used for identifying the multiple and interrelated work practices, and the use of the digital tools. According to Mol (2003)'s multiplicity of reality, we seek to investigate the reality behind the doing, the being, and the ways multiple objects emerge (the design projects)- known as "objects in the making." We further find evidence of these two perspectives (multiplicity and sharedness) from our empirical case in the context of Common Information Space (CIS) (Schmidt and Bannon, 1992; Banon, 2002). CIS seeks to support communication, interaction, and sharing of information to collectively perform work tasks and organize dependencies among the involved actors.

The paper is structured as follows. In the following section, the conceptual perspectives of CIS are described, followed by an introduction of the research context and method for the study. The findings section describes the events, practices, negotiation, and relations in the early-stage design. The discussion section synthesized the conceptual perspective with the empirical data presented in the finding section. This paper then concludes by outlining some limitations and implications for research and practice.

Theoretical framework: common information space (CIS), revisited

The concept of common information space has a particularly long history in the field of CSCW. This is due to the way it was early recognized as foundational to the area. In the inaugural issue of the CSCW journal, the two editors explain the notion as "a central archive of organizational information with some level of 'shared' agreement as to the meaning of this information (locally constructed), despite the marked differences concerning the origins and context of these information items" (Schmidt and Bannon, 1992). In an elaboration of the concept about a decade later, Bannon (2000, p. 3) points out how "In any cooperative work situation, there is a need for some form of communication or information sharing between actors, implicit or explicit, to ascertain what features of the work are of note in that specific situation". Proponents of the concept were clear that absolute or complete sharedness was theoretically futile, hence "The term "common" was used to lessen the connotations associated with the word 'sharing'" (Bannon 2000).

Still, it is reasonable to understand the pioneering of CIS to advocate a position of a *significant* degree of sharedness of meaning, generating a stream of research analyzing how formal and informal articulation of sharedness was established across geographically, professionally, and institutionally heterogeneous settings (Munkvold and Ellingsen, 2007; Bendiksen et al., 2005; Bertelsen and Bødker, 2001).

A second position on CIS developed, insisting that only a *modest* degree of sharedness was necessary, that the former position exaggerated the degree of sharedness required. The formative ideas for this development came from the insights captured by the concept of boundary object, BO (Star and Griesemer, 1989; Star, 2010). Empirically first, then theoretically, scholars unpacked how collaborative practices among diverse communities regularly took place despite significant differences in perceptions, meaning, i.e., lack of shared understanding. The modest requirement of sharedness was the ability across the communities to recognize the identity of the ‘object’ in question. In short, BO opens up to a view with collaborative practices unfolding because of, not despite, lack of shared understanding, with the moderate requirement of agreeing on the identity of the object. The insights of BO stimulated a rich stream of research into but a modest degree of sharedness (Carlile, 2004; Levina and Vaast, 2005; Hepsø, 2009; Fields et al., 2003).

Finally, a third position has developed, arguing that even the modest level of sharedness within a BO-inspired perspective on CIS is overstating the case for sharedness; only a *minimum* degree of sharedness is required. More specifically, there is no (!) sharedness between communities in everyday situations: their practices, vocabulary, and attention differ. As Mol (2003, p. 72) puts it: “It is not a question of looking from different perspectives either as differences are incompatible; there is not one object but multiple; objects are multiple and ‘make a patchwork’”. How, then, is communication at all? Communication, Mol argues, is but a temporal and contingent achievement that solves the immediate need for coordination, synchronization, exchange. Having completed the exchange, the communities go back to the everyday situation of working independently without any degree of sharedness of understanding. Again, such a minimal requirement for sharedness - not so much ‘sharedness’ as loosely coupled, independently working communities occasionally engaging in exchange - has informed a stream of research related to CIS (Oborn et al., 2011; Rolland and Monteiro, 2006).

To summarize, the concept and underlying motivation of CIS have shaped the field of CSCW from its inception. The theoretically intriguing and empirically crucial issue of what degree of sharedness of understanding is required remains contested and unresolved in CSCW, opening to further empirical studies of how, when, and where adequate degree and form of ‘sharedness’ is achieved.

Case context: early-stage design in Engineering Company

The case for our study is an early-stage design of industrial equipment in an Engineering Company (EC)¹, which is a multinational engineering, procurement, and construction (EPC) company delivering equipment to offshore oil and gas production. The equipment EC provides ranges from a single piece of equipment to an entire production facility required for offshore oil and gas production. Companies contracting equipment from EC are mostly oil companies operating offshore fields. EC delivers mainly subsea equipment, that is, equipment for production facilities that sit at the seabed.

Contracting companies contact EC during the early-stage design of the project to study and execute the subsea installment projects. During early-stage design, EC delivers an overall concept for the design of the contracted subsea installation. With limited knowledge of the geological conditions from the customer's initial stage of the project, the phase of early-stage design we focus on is characterized by a high degree of uncertainty about the design and configuration of appropriate technological solutions and methods to be provided by EC.

Therefore, early-stage design project in EC is *open-ended*. The design theme in this context can vary from designing the humongous subsea equipment to pipes and flowlines, reservoirs, wells, and the complete oil field. With so much open in early-stage design, there is a dire need to communicate, discuss and share ideas, concepts, and design sketches among the involved stakeholders in EC-internal and external stakeholders.

An early-stage design team (ESDT) is established in EC to plan and manage the early-stage design work. The group consists of 10-15 professionals, which is again split into 8-10 packages. The team members increase to 30 -35 depending on the subsea and surface-related product lines. The *packages* or *product line* vary, for instance, wellhead, the Christmas tree, the template, manifold, etc., which refer to the various parts of the subsea and surface infrastructure. The packages are assigned with a predefined number. And a specialist from each of these product lines is responsible for their respective packages. Besides, the team also includes technical managers, system engineers, product project managers or lead engineers, project engineers, etc. The crucial point for effective collaboration is maintained through weekly meetings (more often) where different specialists communicate from ESDT meet (physically or digitally) to discuss, review, and agree on the various solutions. This kind of awareness is essential for providing a holistic overview of the entire project, where the different packages are interdependent in terms of requirement handling and designing the technical solution for the successful completion of the entire early-stage project.

¹ EC is a pseudonym used for anonymization purposes

Central to collaboration and communication of work practices is the ability of digital tool and their use by different professionals to work collaboratively by sharing information. We have grouped the digital tools into three categories based on their use during the early-stage design of any project within the team members for various purpose. Some of the collaborative tools are explicitly stated as the customer requirement for EC, such as Skype or mutual SharePoint. For instance, as the Project Manager commented: “(Oil operators) have requirements on which tools to use, how to store things, how to communicate with regards to collaborative digital tools or video conference meeting rooms”.

Considering the prominence of digital tools used in the everyday work of the team members, we have outlined some of the tools used by the design group in the following Table I.

Table I. The list of digital tools used in the early-stage design in EC

Purpose	Type of digital tools	Functionalities
Collaborative	SharePoint	Document storage
	Xaitporter	Documentation tool for multiple user
	Let’s Agree	Interface register for agreements
	Skype	Communication
	MS package (word, excel, PowerPoint) Adobe reader	For creation, documentation and presentation
	Email	Communication and document sharing
Design/Specialist	AutoCAD, NX	2D drawings
	CET	Storage and retrieving old field layout
Project management	RMS	Requirement Management system

Methods

This paper is based on an interpretive case study (Klein and Myers, 1999; Walsham, 1995) of communication and collaboration in an early-stage design of subsea installations within EC. Data was collected through semi-structured

interviews (Kvale and Brinkman, 2009) with professionals involved in early-stage design. The interviews were conducted in two stages. The first stage of data is collected from January 2020 until July 2020, and the second stage is collected from June 2021 until November 2021 (and is continuing). With the basis of a small initial sample of engineers, managers, and developers at EC, we employed snowballing to identify additional interviewees. We have collectively conducted 35 semi-structured interviews within EC and another 13 semi-structured interviews in the contracting company with managers, engineers, and consultants (on a contractual basis). Some of the interviews were face-to-face, while many were online due to the ongoing corona pandemic. The duration of the interviews was approximately 60-90 minutes. The interviews were audio-recorded and transcribed.

We performed thematic analysis (Braun and Clarke, 2012) to analyze our empirical material deductively. Thematic analysis is a method for identifying, analyzing, organizing, describing, and reporting themes found within a data set (Braun & Clarke, 2006). All authors participated in the analysis of the empirical material-interview data. The analysis is also informed by the second and third author's experience from sustained engagement with the oil and gas sector over the past decades.

Our data analysis could be seen unfolding predominantly in three stages. The deductive impulses come from CIS as some of the ideas about the CIS were already present. Secondly, inductively looking at empirical examples. Finally, combining both approaches as part of the analysis process.

We marked the work practices, the use of digital tools, and how different people interpret/share the understanding related to collaborative work to explore the dimensionality of sharedness in our empirical data, particularly the presence of multiplicity. We elaborated insights on the potential use of sharedness in multidisciplinary settings. We built on the key challenges associated with the perspective/lens of "*common or shared information space*," mainly studied to develop and share the knowledge among *heterogeneous groups* (Schmidt and Banon, 1992). The early-stage design phase of EC shares similar cooperative work settings investigated previously by CSCW studies contributing to CIS of establishing and maintaining a sufficiently common understanding of the field of work and the early-stage integration of clients. Our findings are presented in the next section.

Findings: communication and collaboration in the early-stage design of subsea installations

This section focuses on communication and collaboration in the early-stage design of a particular piece of subsea equipment: the Christmas tree (X-mass tree). The X-mass tree is the piece of equipment that connects the subsurface pipelines with the production facilities. It consists of safety valves and other regulatory mechanisms

to control the flow of hydrocarbons from the wells that are drilled hundreds of meters into the seabed. X-mass trees are, at one level, a standardized piece of equipment, but given that “every [oil] well is unique,” as noted in the introduction, X-mass trees always need to be designed and delivered to the conditions of fields and their wells.

To illustrate what constitutes shared understanding and how much of this shared knowledge is required in this complex design setting, we elaborate on an example where collaborative work practices in and across organizations support cooperation in the arena. We follow the trail of our case based on the design of a X-mass tree and the various process (planning, design, review and finalize), tools, and people involved in it. Nevertheless, the people work both as specialists among their respective professions and work in the multidisciplinary design setting- “aligning to both specialist and multidisciplinary practices” (Oborn et al., 2011). The design of such equipment is as complex as the entire project considering the lead facilities (Scotland), research facilities (Houston), and the early-stage design facilities (Norway) to be geographically dispersed. EC received this sub-project of designing the Christmas tree as part of a bigger project as described earlier in the case context.

Initial negotiations: Informal

Delivery projects begin with - communication in the early stage - continuous engagement from the salespeople from EC with the client to look for new or upcoming opportunities. Most of the early communication between the two stakeholders is a mix of phone calls and physical meetings. Typically, the salespeople prefer to have a physical meeting with the client during their early engagement. The informal chats surrounding coffee machine (known as the coffee machine phenomenon although becoming rare with the current mobile “work setting”) provides opportunities for informal and serendipitous communication where the salespeople may receive more information regarding the upcoming projects that pave the path for further coordination between these two stakeholders. Such chats may last for a few seconds or minutes, provide a space for *informal collaboration* and *awareness* of ongoing or upcoming activities (Twidale, 2005). Awareness (in CSCW) is fundamental to coordinate activities and sharing of information, which in turn, are critical to improved collaboration (Schmidt, 2011). As the Director of Engineering put it:

“(...) but maybe we'll get some more information around the coffee, mostly in the more informal and it's also to be in personal relationships, which is very important for the salespeople.”
(Interview excerpt, Director Engineering of EC)

However, the level of required articulation work varies according to the workplace context. As Bossen (2002) put it: “At the wastewater plant, a talk during the coffee breaks is sufficient to coordinate the workers’ different tasks, whereas work at the

hospital ward is dependent on a much higher need for continuous coordination.” Similarly, the sheer coordination activities during the coffee breaks for the co-located early-stage team members are not sufficient for collaboration. But this informal exchange of information leads to more formal and continuous coordination between EC and its customer. Once they agree on an opportunity, a more formal meeting (Teams or face-to-face) is arranged between EC and the contracting company involving the high-level Engineering Managers (managerial decision-making process) for detailed discussion. The *communication* and *negotiation* ensure an agreement ensuring the engagement of EC in the early-stage design of the projects.

Setting up a delivery project: Formalizing the project organization

Once EC and the customer have agreed to proceed with early-stage design, the functional requirement from the customer is received through a written document (mostly a detailed pdf document) and sent by email from the contracting company to EC. Subsequently, the early-stage team from EC starts to study the requirement and engage in the design for that project. The involved team members then assess the requirements by evaluating the economic viability, detailing the technical details and engineering aspects to reduce the future risk related to the project. Through multiple negotiations with the contracting company and detailing of the proposed solution the team (ESDT) then enters the phase where they detailed out the most optimized design solution. Based on the feasibility of the proposed design, EC may or may not receive a formal contract from the contracting companies. Primarily, this phase then deals with multiple negotiations between the different stakeholders and is critical for obtaining the contract by providing the most feasible engineering design solution to the customer.

To enhance proper communication and coordination within the ESDT team inside EC, a *modular working structure* is followed. An ESDT team consists of an early-stage design manager overarching the entire team and followed by the engineering managers, who are responsible for all the technical details. Based on the different product groups, the lead engineers (the specialists for their respective product lines, often known as *packages*) are also part of the design team. At this stage, the involvement of departments is contingent upon the challenges associated with the projects, so is the participation of the professions, disciplines, and various people. The following vignette states the above perspective well regarding the significance and the reliance of the people involved in such complex task:

“(…) So, the point is that from then to pull some of **the older resources needed** for that specific (project) and that’s kind of a cocktail you mix every... for every opportunity. It really depends on what the challenge in that reservoir. Is it plain simple bread and butter, simple tie back in the, in our case the in the Norwegian Continental Shelf? We need maybe half of these actual

resources. If it's a very remote, very technically challenging project, in faraway from us, then that will be a larger involvement from a larger team and that it grows with as the project matures." (Interview excerpt, Field Study Manager)

The Field Study Manager recounted that for every new project, they discuss and agree on "how much is genuinely new; how much can they lean on/use insights/design/solutions/ procedures from the earlier project".

To accomplish collaboration at this stage by achieving and maintaining a common understanding of the complex tasks is challenging due to the lack of integration by various organizational and individual factors. Moreover, the digital artifacts present in the multiple interfaces to share the information is fragmented. For instance, at times when EC has received the functional requirement for designing the subsea equipment, another oil operator might have the wells, other EPC contractors might have flowlines, a third might have risers, a fourth might have the platform, another one might have onshore terminal and another one the export line. This diverse setting indicates the required volume of cooperative activities in these mega projects in distributed, geographic, and disciplinary diversity. According to the feasibility study manager, "*The number of crossings takes time and drives schedule,*" influencing EC's overall front-end study design. So, the question then arises, *how do the design activities occur during the early stage, and how much shared understanding is required during the collaborative work of such a complex design process?*

Planning: Requirement analysis and management

The primary requirement for the project is received from the contracting company and documented in a written pdf report. The high-level requirements from the customer include the details such as pressure and temperature on the reservoir, water depth, flow rates, fluid composition, suboceanic data, drawings, etc. The early-stage design team from EC then translates the project requirements from the client into their "requirement document." This document is part of the Requirement Management System, a partially digitalized, mostly handwritten pdf document, with the details of the customer requirement. And the ESDT makes sure that the requirements are clearly split out for the different sub-systems (packages) like Christmas tree, Control system, Template, Manifold, etc. Then, the Christmas tree group located in Scotland and Houston translates the requirement and starts detailing that further into general arrangements, drawings, specifications for the specific subsea equipment, and make sure that they build up to documentation for that Christmas tree according to the overall requirements.

This ESDT for the subsea equipment consists of 30 to 40 departments with multidisciplinary, geographically distributed professionals. One of the key departments involved in ESDT is the Engineering division of REMS (Research Engineering Management and Supply) with one study manager, technical manager,

Project Manager from Norway with other departments from the construction facility located in Scotland and Center of Excellence at Houston, making the team/group size to 20 approx. The point of departure for the ESDT is to start with the shared requirement document for the subsea equipment (X-mass tree). As of today, EC is using multiple digital tools to support the requirement management system by adopting a more database approach. Yet, ESDT must follow a more *modular* work distribution during the early-stage design.

“And then what we are aiming for this (Requirement Management System), of course, to have it more **database-centric**, so it's one database that everyone can pull information from, and it's all connected. But that is not the fact today. (...) It's so Stone Age, the way we do it today. And, I think, I have been in EC now for 30 years. And I say that maybe we are still working the way we did 30 years ago. So that means it's hasn't moved much.” (Interview excerpt, ESDT Manager)

When it comes to breaking down these requirements into activities and subtasks, the ESTD teams often use the excel-based CTR catalog as standard industry practice. This is a globally agreed digital artifact for work breakdown in the industry.

“(...)is the CTR catalogs we are using in studies, as an Excel-based. It's all based on this work breakdown structure we have internally. So, I think it's actually adopted by the industry as such, you know, but it's globally agreed what number every subsystem has. So, all those CTRs are based on that work break structure. And it's clearly defined the deliverables in each of the packages. So, that's kind of the first breakdown of the tasks, and so, you know exactly what you're supposed to deliver. And you know, also the dependency of that in order to do that delivery to make it, you need to get the input from somewhere, right?” (Interview excerpt, ESDT Manager)

Design: Aligning through meetings

Design work is aligned through a wide array of meetings. These can be design meetings between engineers, more higher-level review meetings with team and project managers (at times includes customers), and strategic meetings involving top-level management.

To regulate their design work, ESDT for the X-mass tree conducts multiple video meetings for the co-located participants and for the remote participants and HACCP (Hazard Analysis and Critical Control Point). As per lean methodology, HACCP principles are used by the oil and gas industry equipment manufacturer to establish a systematic approach to hazard and risk analysis. Design review meetings are primarily essential for the ESDT to align regarding the design. The interdisciplinary group members bring their ideas and understanding related to the flow of hydrocarbons through the subsea tree, the potential hazards, and the failure modes to these meetings. Mostly the ESDT brings 3D/2D models to these design review meetings. The designers then use the model to explain and show the

potential hazards using various design software and computer systems. This provides an overall understanding for the meeting participants. A shift of work practices is observed as the designers bring models rather than written pdf documents to these meetings. The models include both logical flow models and physical design.

“And models, of course, we're also using models a lot. And this drive-in, especially in the high level and lower-level design reviews to use models, instead of written documents, they kind of have a rule or a guideline that we are not checking where the comma should be or if you spelled things correctly, that's not the intent of the design review. So, **where we have succeeded the most, I think, the way used to model-based design review to bring up the models and the designers to use the model and explain and show where the hazards are and understand.** So, everyone involved is **fully aware** of what we're talking about. (...) **I think we have changed the company towards some more related design reviews instead of checking small spelling errors.**” (Interview excerpt, ESDT Manager)

Spatially, the ESDT is distributed across the globe, located in different locations such as Scotland, Houston, and Norway, whereas the different actors involved belong to different specialist groups. At a more conceptual level, the task is divided among the members based on their competencies and work experience. The articulation work or sharing of knowledge across the design team is achieved through multiple coordination meetings (both formal and informal), design review meetings, risk assessment meetings, different artifacts (digital and physical), and peripheral awareness.

One such example is the negotiation process during the design review meetings within the different communities of practice where ESDT and the contracting companies mutually agree upon the design review document or resolve the disagreement based on the peripheral awareness (Bossen, 2002; Bertelsen and Bødker, 2000).

“**(Disagreements/challenges during review meeting)** happens very often, of course, and it's what I call *peripheral engineering*. *People have their preferences based on their experience*. I have been in the industry for 30 years. I do have my preferences. **I know what's working and not and I may not** be that good to come up with arguments why it doesn't work, right. So, but of course, when the client has been through that journey himself, he may kind of have a clearer idea and what's wrong with them, right, instead of listening to the designers and be open for new ideas. Also, sometimes we run into some debates on whose design is the right one, because instead of looking at data from a holistic view and look at the benefits from the different alternatives”. (Interview excerpt, ESDT Manager)

The problems encountered during the change of design in one package must be addressed via proper coordination between the team members during the weekly meetings as the changes may impact the design of another package. As the Technical Manager puts it, “So, you have to have a system overview. (...) If

anything changes on one package, does it have an impact on the other one?" The information-sharing may take place through the help of various design tools, collaborative digital tools, or in multiple meetings at this stage. When different groups of people are involved in such "design information space" with given time, space, the proper *alignment* of the people and technology is essential for carrying out the coordination process within the group.

The customer of EC is also present in some of these design review meetings and provides feedback for the design based on their expertise on exploration, oil field, and seismic data. As the ESDT manager said during the interview, "(...) *Client is very much involved in the design reviews and making sure that they are informed all the time, and then they are actually contributing a lot as well* because they will give input not necessarily has been given as a starting point, but then they can get to. I think that iterative process including the client is extremely important." Once the design review meeting is done, the aggregated document (a shared digital model as of specification, drawings, and schematics) is delivered as a package. So, there is a continuous interaction between the groups (ESDT) while populating the final deliverables through XaitPorter (cloud-based tool for document production). The digital tool support for the final deliverable may vary from project to project. The interaction within the team may continue with the help of Microsoft Team's chat functionality to stay in the same communication spectrum to avoid any deviation in the deliverables.

The ESDT follows the iObeya tool founded on Lean and Agile principles for empowering the distributed team and keeping the communication open within the team. But often, it is difficult to negotiate and bring all the stakeholders on board during the design work, which may lead to future challenges in execution. To avoid that, a high-level technical decision meeting or strategy meeting is arranged to evaluate the challenges and risks. Often the top-level management is present with the ESDT to assess the risk and uniqueness associated with the activity. The ESDT manager's comment illustrated the challenges (frustration) and the acceptance of those challenges towards a working solution during the design task.

"(...*Internal communication*) is good with iObeya and all that. There are a number of tools and how to do that in a good way and break down tasks in manageable pieces. (...) But in this version, (there) is a lot of fixed and flexibles that you need to consider something you're not able to move due to **political stuff internally and something new, you're able to change and configure it in a different way**. That iterative process is not often that easy to get everyone on board. And we do have some success now with something called **technical decision or strategy meeting**. (...) where we bring the vice presidents and top-level management together. We talk about the challenges within the project, any uniqueness in the project, and we risk it, if it's something that needs to be elevated to a higher level, or if it can be continued?"(Interview excerpt, ESDT Manager)

Information sources and use of collaborative tools

The various design tools used, interactive dashboards, the design review reports contribute towards the shared understanding between the designers, lead engineers, project managers, and the customer involved in the design process.

“We often shared it in documents as such. But we can also, we have an interface tool, we call it Lets Agree where we share a lot of (information). It's a handy tool for us, we also use the client tools like PIMS to share documents, etc.”(Interview excerpt, Technical Manager)

As is evident, relevant information related to design is spread across different systems and formats (including paper and pdf), making it difficult for the team to access correct information. EC has designed and developed another tool, CET (collaborative engineering tool, the name changed for anonymization purpose), to avoid the shortcomings of existing tools and processes. This is recently added to the already existing set of early-stage design tools, which can work as a platform where multiple specialists can work together in a multidisciplinary setting to access and share past design layouts. Moreover, this tool provides an overview of the current work processes surrounding the early-stage design. As CET holds much sensitive information related to cost, schedule, and perpetrated knowledge about the early-design work, it is essential to share the understanding between the involved stakeholders. To overcome the challenges related to information sharing, EC provides limited access to specific users of CET. The transition to CET aims at significantly changing the task and role of the various professionals involved, thus redefining the collaborative work processes. This tool acts as a common artifact and provides the medium for different specialists to hold a modest degree of understanding to understand each other's coordinated design activity which is reflected from the comment of the Vice President and Head of Subsea.

“So, what we want (during early-stage design), is that we would like to see the rig-side with the client. You are the drilling department, I am the CET, (stakeholder). We see together and mature different value scenarios, and some time different technology scenarios together, and we get that iteration loop. Because what you do in Subsurface affects how we design our subsea, how we drill the wells. So, all this hangs together as an optimization wheel. And as we mature, you get more and more equations and, less and less unknowns.” (Interview excerpt, Vice President, and head of Subsea)

To conclude, in an ideal scenario, the designers from different packages work towards the design solution. However, the politics of workspace practices may vary in customer communication and managerial decision-making, creating communication issues within the multidisciplinary team requiring a change of practices. As put by the technical manager during the interview:

“(…) So, what happened in a lower-level, lead to lead meeting with a client, **is maybe not shared** on a high level, and it's a crosslink with us and the client that is making the communication a bit harder. So that's kind of the startup communication issues we have now, of course, then we need to align to make sure that we tell this setup for making sure the communication lines are working better.” (Interview excerpt, Technical Manager)

The above example highlights that the cooperative work setting is different in diverse organizations. EC and the contracting company possess other challenges during the collaborative design meetings. It is challenging for the technical manager to ensure that the 30/40 product lines or packages are well informed and talking to each other. However, due to different work settings present within the client, a lower-level meeting with lead design engineers and the client is not communicated to the higher level. Thus, it creates a communication gap between the different stakeholders, which needs a change of work practices.

So far, we have presented some of the results showing the process of communication and collaboration in early-stage design. We have found the trail and traces of the various communication processes mediated by human and digital technology during design. Some of the stakeholders - human actors (as referred by Banon and Brødker, 2002) physically share the workplace and are co-located, whereas some of the other design team members are in a remote location – present in a distributed work setting. Evidently, information is spread over multiple digital and physical artifacts (databases, files, folders, reports, design drawings, models, prototypes, digital tools, and so on.) but is difficult to access. Moreover, information sharing among design team members is limited and restricted due to the highly sensitive design data.

Discussions

The different work settings influence the team member's understanding or perception of the design context. Whereas it is easier for the co-located team to make sense of the shared context, it is difficult for the other team members to understand the same in a distributed work setting. The role of boundary objects and boundary spanners in terms of artifacts and human mediators is found in our X-mass tree design work context. We also found strong evidence of the third perspective of shared information spaces influenced by Mol's multiplicity, apart from CIS's two much-studied aspects. Contrary to the notion that the communication process aspires to collaborate by effectively sharing required information, we found that most design activities, i.e.-vocabulary, understanding, tools, practices, and actors, are different. It is not the mere understanding or interpretation of the activities which is different in the design stage, but everything is different. Thus, our finding supports Mol's 2003 "multiple" concepts. The goal at this stage is not to accomplish a complete shared understanding (information

sharing) as described in the CSCW literature within the stakeholders but to achieve collaboration with the least possible or minimum shared knowledge to accomplish the design task.

As outlined in our theoretical framing, we distinguish three *degrees* of sharedness in the CIS literature. Based on the findings, we identify and discuss these three different versions of CIS, namely *significant* (CIS1; Munkvold and Ellingsen, 2007; Beretelsen and Brødker, 2001), *modest* (CIS2; Levina and Vast, 2005; Fields et al., 2003) and *minimum* degree of sharedness (CIS3; Mol, 2003; Oborn et al., 2011) as presented in Table II.

Table II Types of Common Information Space

Levels	Degree of sharedness	Examples
CIS1	Significant (shared, common)	Work practices, digital tools, review meetings, participants (ESDT member)
CIS2	Modest (Boundary objects, boundary spanning)	The role of the of the designer in the design review meetings as boundary spanners and the database, design standards, design drawings, models and schematics related to CT as Boundary Objects
CIS3	Minimum (Multiple objects)	Multiplicity of technology use, diversity of work practices and design context, multiplicity of perspectives

CIS1: CIS1 corresponds most closely with early elaborations of CIS that focuses on sharedness and common understanding (Schmidt and Banon, 1992; Bannon and Brødker, 1997; Bannon 2000; Munkvold and Ellingsen, 2007). We find this perspective in the context of the early-stage design of X-mass trees context which is heterogeneous with respect to space, artifacts, and actors. The “shared” meaning presented in most of the foundational CIS work assumed a significant degree of sharedness (knowledge work) across the different communities-in our case different disciplines. Schmidt and Bannon (1992) conceptualized the very early CIS as:

“how people in a distributed setting can work cooperatively in a common information space - i.e. by maintaining a central archive of organizational information with some level of ‘shared’ agreement as to the meaning of this information (locally constructed), despite the marked differences concerning the origins and context of these information items. The space is constituted and maintained by different actors employing different conceptualizations and multiple decision making strategies, supported by technology.”

The concept of CIS was further expanded and explored by Bannon and Bødker (1997), Bossen (2002). It was designed to explore the then-existing notion about the role of technology and sharing information in collaborative work. It was further developed by Bosses (2002) to help the researchers have a common understanding of a work situation/context and its coordination by embracing the combination of information, its representation, and interpretation. In the context of CIS, information sharing is never without a problem. Although the information is put in common, how the team members contextualize the information according to their work need and interpret/use the information is important. Therefore, it is relevant to look at how CIS is incorporated into the everyday work practices of an ESDT.

From the findings section, it is evident that information is “shared” in documents such as design reports and through multiple tools. The information is further contextualized and decontextualized by the different team members such as the designers, project managers, lead engineers, or the customers. Although the actors involved interpret (practical understanding) the information differently, they work towards a final design solution.

CIS2: This version of CIS in CSCW literature addresses the research related to a moderate degree of sharedness of understanding, information in multidisciplinary collaborative work where the tasks are interdependent, and the existing dependencies must be addressed for achieving the desired/shared goal. As defined by (Banon and Brødker, 1997), thus, the CIS is conceptualized as malleable and open in nature introducing the boundary concept.

“On the other hand, in distributed work settings, there is a much greater need for refining and “packaging” information into a meaningful context, in order to maximize the likelihood that the intent of the message is received appropriately, and the recipient is also required to expend some effort in order to “unpack” this information, and hopefully be able to re-create the context of its transmission.”

In the above scenario, boundary objects act as the carrier of critical information across the various communities of practice. As defined by (Star and Griesemer’s (1989):

...objects which both inhabit several intersecting social worlds and satisfy the informational requirements of each of them. Boundary objects are objects which are plastic enough to adapt to

local needs and the constraints of the several parties employing them, yet robust enough to maintain a common identity across sites. (p. 393)

Accordingly, the various design drawings, models (diagrams, schematics) brought to the design review meetings act as a “*common object*” across different professionals (boundaries) present in the design review meeting. In the given context, the physical or digital artifacts (design drawings, prototypes, tools such as CET) signifies a certain level of overlap (*moderate degree of sharedness*) in the shared meaning related to the X-mass tree design between the different professional communities such as customer, technical manager, project or study manager, designers, and customer- thus facilitating inter-, intra-organizational knowledge transfer.

Furthermore, the design review meetings represent a distributed work setting concerning the time and space boundaries and geographically distributed team. According to Banon and Brødker (1997), the mechanism used to “holding in common” the information varies accordingly during the design review meetings. As opposed to the co-located team and physically shared workspace, the distributed teams as in design review meetings, the designers must translate the design knowledge into a meaningful package to create a shared understanding among all the participants (different professional groups) in the review meeting, acting as the *boundary spanner*. The recipient in this context ‘unpacks’ the information and interprets the work context.

CIS3: This third thematic area and last version of CIS in CSCW literature are shaped by Mol (2003)’s multiplicity perspective - hence a minimum or no sharedness of understanding is required while performing the design-related work practices using the various design tools.

Though multiple studies have explored the phenomenon of CIS from the significant and modest level of sharedness, very few have explored the minimum degree of sharedness in understanding the CIS. Our study thus adds to the limited stock of minimum sharedness. As AnneMarie Mol (2003) puts it, the same disease of atherosclerosis is ‘done’ – or is – differently, our case illustrates how the perceptions, practices differ. However, these differences of need must be resolved temporarily and contingently to reach decisions on treating the patient who has atherosclerosis. Thus, the third perspective of sharedness, which we are interested in, deals with the multiplicity of perspectives in the collaborative work setting. As discussed by Mol, multiplicity assumes that there is no *one reality*, but *multiple realities* co-exist as in the everyday hum of activities, vocabulary, tools, practices, and understanding differ. Hence the way to understand the multiplicity is to understand how the objects act in those multiple practices and realities as we infer objects understanding from their practices. Mol (2003)’s ethnographic study of the disease (atherosclerosis) runs like this on multiplicity:

If practices are foregrounded there is no longer a single passive object in the middle, waiting to be seen from the point of view of seemingly endless series of perspectives. Instead, objects come into being – and disappear – with the practices in which they are manipulated. And since the object of manipulation tends to differ from one practice to another, reality **multiplies**. The body, the patient, the disease, the doctor, the technician, the technology: all of these are more than one. More than singular. This begs the question of how they are **related**. For even if objects differ from one practice to another, there are relations between these practices. Thus, far from necessarily falling into **fragments**, multiple objects tend to hang together somehow. Attending to the multiplicity of reality opens up the possibility of studying this remarkable achievement. (Mol, 2003, p.5)

Further describing the disease and the multiple associated practices in a hospital setting, Mol says, ‘The body multiple is not fragmented. Even if it is multiple, it also hangs together’ (ibid.: 55). Therefore, multiplicity does not imply fragmentation; instead refers to interlinked practices with limited or no sharedness in understanding how each object works. As conceptualized by Mol, translation is then required to connect between the diverse work practices and support knowledge sharing between the team members. However, translations are never smooth as they can work together and act differently simultaneously (Mol, 2003, p.72).

For instance, the object in our case is a design project which, on the one hand, is brought into being a financial object (considering cost, revenue, and time schedule) in the operator’s (contracting companies) perspective. On the other hand, the project is reframed as an engineering object (task) from EC’s perspective. These two realities, where a project is considered an object but brought into being different work practices and digital tools. And they are translated and brought into alignment at the reporting point. Unlike the boundary objects, translation can be studied using various computer-supported design tools such as AutoCAD, CAM, CET, etc. Yet, these tools are used by the designers and the other group members differently. The customers and the project managers interpret the designers’ drawings, models, and reports in the design review meetings to provide their respective feedback to the design team, negotiate, and make the final decision that resembles the multiplicity of translation activities. Based on the openness and complexity of each design task, how the designers are breaking down the requirement specification and managing requirement analysis before meeting the customer requirement is an essential aspect of translation practices.

Thus, a shift of work practices (multiplicity) is observed in the early-stage design work. Although it is unclear how the different disciplines interpret the different activities in practice, how the managers understand the design task, these differences must indeed be reconciled to some extent to get the final design work done. So, studying the performances (through work practices, digital tools) of the early-stage design work reveals their multiplicity. We interpret this example as the minimum degree of shared understanding required for the CIS3.

In sum, our findings find evidence of all three versions or degrees of sharedness in CIS, ranging from significant via moderate to but minimal. Relative to the CSCW discourse on CIS, however, the most interesting aspect of our analysis is the presence of the third (“minimal”) overlap. This represents an under-researched version of CIS against the dominance of moderate (and significant) degrees of overlap (Mol, 2003, Oborn et al., 2011).

Conclusion

To understand the complexity of the collaborative design work, we have studied the different work practices in an offshore oil and gas industry- in the context of the subsea equipment design. The findings from the empirics show how the various specialists with diverse domain knowledge from EC perform the work associated with design, use different technology to communicate and collaborate across the different stakeholders, decide, and conclude the final design solutions, and at the same time, keeping a close collaboration with the client.

However, it is rather hard to discern the degree of sharedness in multidisciplinary cooperative work. While much CSCW research is devoted to the common or sharedness and boundary objects to study collaboration and information practices, only a few have investigated the minimum degree of sharedness (Oborn et al., 2011; Bjørn and Hertzum, 2011; Büscher et al., 2009). Thus, to put it provocatively, the case for sharedness in collaboration in CSCW has been exaggerated. We show here with this paper that sharedness resides along the continuum of two extremes, i.e., from complete to no sharedness. Usually, the researchers focus on the significant and modest degree of sharedness. But we also then show the extent to which this Mol (2003) influenced multiplicity of practices (with no or minimum sharedness) are present in our empirical case - empirical contribution.

Finally, we would like to point out some limitations when relying on sensitive design data related to large-scale projects. The access to the field observation and relevant interviewees from different locations was challenging due to the COVID restriction, which was one of the major reasons for the very limited number of interviews in the second stage of data collection. The design of industrial equipment requires coordination between different domain expertise, different local context, forms of knowledge, people, tools, and artifacts, which we clearly do not have a complete overview of. We encourage supplementary studies of complex collaboration to explore the degree of sharedness, broaden our understanding of multiplicity of practices inherent in everything, every object. Thus, future research can investigate a minimum understanding of collaborative design work.

Acknowledgments

This research is a part of BRU21 – NTNU Research and Innovation Program on Digital and Automation Solutions for the Oil and Gas Industry (www.ntnu.edu/bru21).

References

- Bannon, L. (2000) Understanding Common Information Spaces in CSCW. Paper presented at the Workshop on Cooperative Organisation of Common Information Spaces. Technical University of Denmark, <http://dmm.cti.dtu.dk/position/bannon.pdf> (last accessed January 2022)
- Bertelsen, O. W., & Bødker, S. (2001). Cooperation in massively distributed information spaces. In ECSCW 2001 (pp. 1-17). Springer, Dordrecht.
- Bjørn, P., & Hertzum, M. (2011). Artefactual multiplicity: A study of emergency-department whiteboards. *Computer Supported Cooperative Work (CSCW)*, 20(1), 93-121.
- Bossen, C. (2002, November). The parameters of common information spaces: The heterogeneity of cooperative work at a hospital ward. In Proceedings of the 2002 ACM conference on Computer supported cooperative work (pp. 176-185).
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101.
- Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 57–71). American Psychological Association. <https://doi.org/10.1037/13620-004>
- Büscher, M., O’Neill, J., & Rooksby, J. (2009). Designing for diagnosing: Introduction to the special issue on diagnostic work. *Computer Supported Cooperative Work (CSCW)*, 18(2), 109-128
- Carlile, P.R., 2004. Transferring, translating, and transforming: An integrative framework for managing knowledge across boundaries. *Organization science*, 15(5), pp.555-568.
- Ellingsen, G. and Monteiro, E., 2003. A patchwork planet integration and cooperation in hospitals. *Computer Supported Cooperative Work (CSCW)*, 12(1), pp.71-95.
- Fields, B., Amaldi, P., & Tassi, A. (2005). Representing collaborative work: the airport as common information space. *Cognition, Technology & Work*, 7(2), 119-133.
- Flyvbjerg, B., 2006. Five misunderstandings about case-study research. *Qualitative inquiry*, 12(2), pp.219-245.

- Hepsø, V. (2009). Common information spaces in knowledge-intensive work: Representation and negotiation of meaning in computer-supported collaboration rooms. In *Handbook of research on knowledge-intensive organizations* (pp. 279-294). IGI Global.
- Hildreth, P. M., & Kimble, C. (Eds.). (2004). *Knowledge networks: Innovation through communities of practice*. Igi Global.
- Klein, H. and Myers, M. (1999) A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems, *MIS Quarterly*, and 23(1): 67-93.
- Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing*. sage.
- Levina, N., & Vaast, E. (2005). The emergence of boundary spanning competence in practice: Implications for implementation and use of information systems. *MIS quarterly*, 335-363.
- Mol, A.M. (2003) *The Body Multiple: Ontology in Medical Practice*, Duke University Press.
- Munkvold, G., & Ellingsen, G. (2007). Common Information Spaces along the illness trajectories of chronic patients. In *ECSCW 2007* (pp. 291-310). Springer, London.
- National Commission 2011. *Deep Water: The Gulf Oil Disaster and the Future of Offshore Drilling*. Nashville, TN: Perseus Distribution Digital.
- Oborn, E., Barrett, M.I., & Davidson, E.J. (2011). Unity in Diversity: Electronic Patient Record Use in Multidisciplinary Practice. *Inf. Syst. Res.*, 22, 547-564.
- Rolland, K. H., Hepsø, V., & Monteiro, E. (2006, November). Conceptualizing common information spaces across heterogeneous contexts: mutable mobiles and side-effects of integration. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work* (pp. 493-500).
- Schmidt, K. and Bannon, L (1992). Taking CSCW Seriously – Supporting Articulation Work. *Computer Supported Cooperative Work (CSCW)* 1, 7-40.
- Schmidt, Kjeld. (2011). *Cooperative Work and Coordinative Practices - Contributions to the Conceptual Foundations of Computer-Supported Cooperative Work (CSCW)*. Heidelberg, Germany: Springer-Verlag.
- Schmidt, K. and Bannon, L.(2013). Constructing CSCW: The first quarter century. *Computer supported cooperative work (CSCW)*, 22(4-6), pp.345-372.
- Star, S.L. and Griesemer, J.R., 1989. Institutional ecology, translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social studies of science*, 19(3), pp.387-420.

- Star, L. S.2 (2010). This is not a boundary object: Reflections on the origin of a concept. *Science, Technology, & Human Values*, 35(5), pp.601-617
- Twidale, M. B. (2005). Over the shoulder learning: supporting brief informal learning. *Computer supported cooperative work (CSCW)*, 14(6), 505-547.
- Walsham, G (1995) Interpretive case studies in IS research: nature and method, *European Journal of Information Systems*, 4(2): 74-81.
- Wittgenstein, L. (2010). *Philosophical investigations*. John Wiley & Sons.

Marref, R., Abou amsha, K., & Lewkowicz, M. (2022): ECSCW 2022 Exploratory Papers Instructions. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_ep13

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_ep13

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

Coupling documentation and communication practices to support integrated care pathways

Rahma Marref, Khuloud Abou Amsha, Myriam Lewkowicz

LIST3N/Tech-CICO– Troyes University of Technology

rahma.marref@utt.fr, khuloud.abou_amsha@utt.fr, myriam.lewkowicz@utt.fr

Abstract. Creating an overview of a patient's situation is essential for medical work, particularly for care actors participating in integrated care pathways. While we notice the failure of national initiatives to create systems to support care actors' practices, our research focuses on investigating the actual practices held by care actors to build these overviews. Our study shows that care actors document their patients' cases and communicate about them. We, therefore, imply that systems supporting integrated care pathways should couple documents and discussions and permit their visualization according to three dimensions: the problems that occur, the source of information (the care actor), and the time.

Introduction

With the increase of multimorbidity, patients are treated by different specialized and non-specialized care actors, often working in different organizations, and interacting within multiple prolonged illness trajectories (Mønsted et al., 2011). This form of care is known as integrated care (Kodner & Spreeuwenberg, 2002).

The OECD countries define an integrated care pathway (Fukuda et al., 2015), also known by The European Pathway Association (E-P-A), as “a complex intervention for the mutual decision making and organization of care for a well-defined group of patients during a well-defined period” (Vanhaecht et al., 2010). The overall public health objective is to define standardized care pathways for chronic diseases such as diabetes (defining the different care steps) in order to avoid the redundancy of medical services (Fukuda et al., 2015).

In this context, many governments started national-wide initiatives to create IT systems to support access to dispersed patients’ information. For example, some initiatives have opted for National Electronic health records (Cresswell, Worth, et al., 2012) that offer a centralized database to store the different patients’ information. However, studies have revealed that national health records are incomplete and outdated (Cresswell, Robertson, et al., 2012) and tend to be challenging to read due to information overload (Amir et al., 2015). Moreover, a lot of information shared in the medical record is context-dependent (Reddy et al., 2001; Zhang et al., 2017). Thus, they risk losing their meaning while traveling from one setting to another.

Other initiatives proposed national systems to facilitate the exchanges among the different stakeholders taking care of a patient (Marref et al., 2022). However, systems that only focus on supporting communication are hard to promote as care actors already use a diverse and evolving set of artifacts to communicate (Fitzpatrick & Ellingsen, 2013).

Facing these limitations, this paper aims to understand how care actors dealing with patients involved in an integrated care pathway manage to get an overview of a patient’s situation. Previous work has shown that care actors working in the same care setting engage in multiple activities for managing a patient’s situation (Bossen & Jensen, 2014), from collecting medical data (from different sources), to communicating informally (on the phone, for instance) when needed. Likewise, care actors in multidisciplinary teams outside the organizational boundaries rely on the shared information inscribed in different documents and communication to build their understanding (Abou Amsha & Lewkowicz, 2016). Therefore, our study examines how the care actors, working in different care settings, use documents and communication to construct, update and maintain an overview of the patient situation. Moreover, we argue that technology designed to support these activities should consider enabling the collection and the visualization of the medical information while also supporting communication (that cannot be scheduled in advance in the care pathway). Thus, we suggest that communications among care actors represent mandatory landmarks for creating a shared medical record supporting their practices.

The remaining parts of this paper are structured as follows: We first review the literature to investigate the role of information sharing and communication in medical practices. Then, we present our case study related to integrated care in the Aube department (North-East of France), investigating how care actors manage to get an overview of the cases they manage. We finally discuss how to link information sharing and communication to support the cooperation among care actors within integrated care pathways.

Related work

Care actors must understand patients' history and present situations to apply their medical knowledge and deliver care (Mønsted, 2015). Following a patient's trajectory permits the actors to walk through the succession of the medical events and the critical and decisive moments in the patient's care journey (Pescosolido, 2013).

In other words, care actors have to get an overview of their patients' trajectory to make sense of lifelong medical information. These overviews are formed of a series of acts, incidents, states, and experiences bound together to build a worthwhile storyline (Feldman et al., 2004), allowing the contextualization of different information (Lutters & Ackerman, 2002). To construct and update these overviews, care actors collect information from diverse sources (Mønsted et al., 2011) and communicate with different care actors to enhance their understanding of the patients' cases (Jensen & Bossen, 2016).

Sharing information in medical practices

Medical records are commonplace to store the corpus of patient information (Berg, 1996) produced and exchanged by care actors to follow the patient's trajectory (Embi et al., 2013). With the growth of IT in medical care, the digitized version of medical records has become a widespread artifact in different settings.

Medical records encompass a list of unstructured, descriptive texts and structured coded documents (Hayrinen et al., 2008) that can be organized into three categories: *the core parts* storing the patient's medical history, the examination notes, the medication sheets, and the physician's progress notes, *the procedural reports* including the results of radiology, blood lab results, etc.; and *the episode summaries* referring to discharge and referral summaries (H. J. Tange et al., 1997).

Electronic health records (EHR) can also be categorized into three families: 1) *Source-oriented records*, organized by the source of information (ex. physicians' entry, laboratories results), presented with the time the information was collected (H. Tange, 1996); 2) *Problem-oriented records*, containing a predefined set of problems that the physicians mobilize to guide their decision-making process (Weed, 1968); 3) *Time-oriented records*, in which medical data is

organized chronologically (H. J. Tange et al., 1997). Nowadays, most of the EHR are source-oriented or time-oriented. Indeed, problem-oriented records did not persist due to the lack of flexibility obliging the users to follow predefined rules of action (McDonald & Barnett, 1990).

Whichever their form, medical records seek to respond to the need to organize the patients' care, document their treatments, and estimate the culmination of their trajectory (Clynch & Kellett, 2015). Some EHRs offer visualization features, permitting navigation and understanding of a large amount of information that would be difficult to make sense of without a proper visualization (Rind et al., 2013). The most common visualization techniques are *lifeline visualizations* (Plaisant et al., 1996) which allow temporal navigation through the medical events and *graphical summaries of patients' status* (Powsner & Tufte, 1994) that permit following the evolution of the numerical medical data registered in the EHR.

Furthermore, medical records are often promoted as tools to ensure the coordination of the medical activities (Berg, 1999), as they offer a shared information space that connects the different healthcare actors (Reddy et al., 2001). However, the information that is generated and shared is context dependent. Thus, the literature debated the possible loss of the information meaning and connotation when moving from one setting to another. For instance, the care actors at a hospital might follow local standards in documenting their interventions that might be difficult to understand outside the hospital. Therefore, the care actors need to de-contextualize the information before re-contextualizing it within the context of the new setting (Zhang et al., 2017). Moreover, it is difficult to identify the relevant information or documents to build the needed overview from the large amount of information stored in the medical records (Mønsted et al., 2011). To face this challenge, (Richter et al., 2016) argued for the necessity to complement 'lifetime electronic record' with communication systems to enable the actors to discuss in order to get a thorough overview of a patient that would favor their understanding of their health situation. However, the research on connecting the information available in the integrated electronic medical records and the communication systems that aim to contextualize them is still missing.

Communication in medical practices

Communication among care actors has a prominent role in their daily work (Coiera, 2000; Gurses & Xiao, 2006; Jensen & Bossen, 2016). Care actors communicate to build a shared understanding of patients' cases (Chao, 2016) and make sense of the different events occurring during the patient's care journey (Paul & Reddy, 2010).

Previous work has identified moments in which communication is critical. For instance, one typical moment is the multidisciplinary round discussions (Gurses & Xiao, 2006). The care actors collaborate to sharpen their collective understanding of each discussed case to plan future treatments and define each care actor's actions. Thus, the patient's trajectory flow results from the constant negotiation between the care actors to make the right decisions (Crocker et al., 2012). Besides, the

continuous communication and sharing among the actors prevent coordination breakdowns that may result in errors in the treatment (Abraham & Reddy, 2013).

While communication occurs efficiently in organizational settings (hospitals, care homes) during team meetings, round discussions, and other moments, it becomes more challenging to support communication in integrated care situations. Indeed, care actors belonging to different organizations face many challenges in contacting each other and keeping the dispersed participants informed about a patient's evolution (Gulmans et al., 2009). In integrated care, the communication and cooperation among care actors occur at moments based on the problem at hand. (Abou amsha et al., 2020) argue that the cooperation occurs in temporary *knots* with evolving care actors depending on the evolution of a patient's situation. Care actors working together to manage a patient at home rely on asynchronous communication (supported by liaison notebooks, in particular) to pass information from one care actor to another (Abou Amsha & Lewkowicz, 2016). However, the vague definition of roles, the lack of transparency, and the hesitation in reaching out to other actors with different backgrounds remain challenges that hinder the communication among care actors (Gulmans et al., 2009; Abou Amsha & Lewkowicz, 2018).

At the same time, patients take a substantial role in the creation of their medical stories as they are the ones who can depict their main ailments (Mønsted et al., 2011). Therefore, the communication with a patient provides the physicians with more insight into the patient's issue and enables them to clarify and reshape their understanding of the situation (ibid). Moreover, in the integrated care context, where the exchange between the actors is an arduous task, the patient is considered the messenger and the coordinator of the different care actors.

Care actors rely on shared medical information as building blocks to construct an overview of a case. At the same time, they lean on communication as a complementary way to enhance and reshape their understanding and, therefore, the overview that serves their care work. We then argue that articulating the information collected during communication might support the overview construction and the overall cooperation among actors involved in integrated care.

Case study: Integrated care in the Aube department

Context

We conducted a study in the Aube department, an agglomeration of approximately 160 000 inhabitants, with eight healthcare establishments: the hospitals of "Champagne Sud", four private clinics, and three medical homes. Besides, the department counts approximately 1300 care actors with private practices (Table I

gives a snapshot of the different profiles¹). The Hospital Center of Troyes (CHT) plays an essential role, as several specialties can be found only in the hospital (e.g., dermatology, ambulatory medicine, and nephrology).

Table I. The distribution of the care actors with a private practice in the Aube department

Medical role	Number of care actors
General practitioners	200
Specialists	147
Dentists	141
Midwives	15
Pharmacists	136
Physiotherapists	236
Nurses	357
Podiatrists	73
Total	1305

Furthermore, the Aube department belongs to Aube-Sézannais, one of the five volunteer territories participating in the French Hospital Federation (FHF) trial of the population-based approach for two populations: diabetics and heart-failure patients. The population-based approach uses the clinical integration model, which that identify target populations, and all stakeholders concerned with their treatment and then develops shared clinical programs aimed at the triple objective of better health and better care at the best cost for this population². Thus, the regional health agency (ARS) worked with a group of care actors to define the different integrated care pathways centered within the perimeter of the department, to respond to the health needs of the department. In this work, the general practitioners, who serve as gatekeepers, have a central role in the care group around a patient.

To facilitate the organization of those care pathways and promote collaboration among the care actors, the ARS launched a project to deploy an information system (named Parcéo). The system, which is currently being customized, aims to provide a shared space enabling asynchronous discussions among actors working within organized care collectives to fulfill the care pathways goals.

¹https://www.grandest.ars.sante.fr/system/files/201705/PRS2_Etat_des_lieux_07_RESSOURCES_HUMAINES_SANTE_20170519.pdf

² <https://www.fhf.fr/Offre-de-soins-Qualite/Systemes-d-information/Mettre-en-oeuvre-la-responsabilite-populationnelle-la-feuille-de-route-de-la-FHF-en-matiere-de-systemes-d-informations#:~:text=La%20Responsabilit%C3%A9%20populationnelle%20repose%20sur,de%20sant%C3%A9%20d'un%20territoire.>

Method

To understand how care actors cooperate in this context, we conducted a qualitative study focusing on care actors participating in the care pathway of diabetic patients.

We recruited care actors in Troyes, both with private practice and employed by care organizations. We sent over sixty invitation emails followed by phone calls to present our study. We received 24 positive answers, but we succeeded in planning only 20 semi-structured interviews between January 2021 and September 2021. The interviews lasted 45 minutes on average, and the interviewees' experiences varied between 6 months and 40 years (see table II). Due to the third and fourth waves of the Covid-19 pandemic, we rescheduled several interviews many times, and six interviews were conducted online through videoconferencing systems (Zoom and Teams). We interviewed seven care actors with exclusively private practice, three with mixed practice (private and employed), and ten care actors employed by eight different care organizations in four different cities of the department (see figure 1). The interviews aimed to investigate how care actors build and maintain an overview of their patients. All the interviews were audio-recorded, transcribed, and analyzed using open coding. We iteratively coded the data; in the first round, we identified the artifacts used during the consultations to document the patient's information, procedures, and prescribed treatments. Then, we analyzed the information exchanged among the different care actors and their role in constructing the case's overview. Finally, we focused on care actors' strategies to complete their medical narrative.

Table II. Data collected

Id	Position	Workplace
Endocrinologist_1	Endocrinologist	Troyes French mutuality health center
Endocrinologist_2	Endocrinologist and head of the service	Troyes hospital Center
Endocrinologist_3	Endocrinologist	Liberal activity
Dietician_1	Dietician	Troyes French mutuality health center
Dietician_2	Nutritionist dietician	Between the private clinic of Champagne, the Hospital Center of Romilly-sur-Seine, and a liberal activity
Dietician_3	Nutritionist dietician	Between the private clinic of Champagne, the ANPAA ³ , and a liberal activity

³ The National Association for the Prevention of Alcoholism and Addiction

GP_1	General practitioner	Liberal activity
Endocrinologist_4	Endocrinologist	Liberal activity
Nurse_1	Therapeutic education Nurse	Troyes French mutuality health center
GP_2	General practitioner – day hospital	The hospital center of Romilly-Sur-Seine
Podiatrist_1	Pedicure podiatrist	Between the Hospital Center of Troyes and a liberal activity
GP_3	General practitioner	Between Troyes and Brienne-le-Château Public Mental Health Institutions
Nurse_2	Therapeutic education Nurse	Troyes hospital Center
Nurse_3	Nurse	Liberal activity
Podiatrist_2	Pedicure podiatrist	Liberal activity
Lab_Doctor	Laboratory doctor	Private medical analysis laboratory
Endocrinologist_5	Endocrinologist	Troyes hospital Center
GP_4	General practitioner	SOS médecins
Kenitherapist	Kenitherapist	Liberal activity
Pharmacist	Pharmacist	Liberal activity



Figure 1. The map of the cities targeted in the field study

Results

A patient's care pathway engages multiple care actors. Most of the time, the patient's GP triggers the integration of the patient in a pathway. Once in the pathway, the patient moves from one care actor to another using the referral letters or prescriptions. In general, patients carry the medical documents they receive during care episodes (e.g., consultation summaries, lab results, and discharge reports).

To get an overview of a patient's trajectory, care actors rely on different sources: the documents and interactions with the patients, their relatives, and, more importantly, the other care actors involved in the patient trajectory (including paramedics and social actors). Based on this global understanding, care actors then write summaries for their personal or other care actors' use.

Getting a case overview

After a patient's visit, care actors write summaries in their local medical record system (digitalized or paper-based records). To do so, they start seeking information about the patient's situation. When care actors see patients for the first time, with no updated or shared medical record, they rely on the patient to answer their questions and as the vehicle of the different documents, as a diabetologist put it:

"The problem is to have a medical record that is not only kept by the patient... That is to say, the DMP [the French national medical record] has been launched, but the problem is that it takes time to fill in the data, and then we do not have the tools to open them when the patient is here in consultation." **Endocrinologist_1**

However, if the patient or their family members cannot give precise answers and provide all medical documents, care actors need to communicate with the other actors around the patient. Therefore, they reach out to the different correspondents, starting with the patient's GP, seeking information and requesting documents as the GP working in the mental hospital put it:

"Our patients are generally unable to tell their stories [...], the reference documents remain with the specialists because they need to focus on treating the psychiatric pathology. Therefore, when it came for us [the general practitioner of the mental hospital] to have more information, we call the emergency department, and we call back the patient's attending physician; when we manage to have them on the phone, we usually call the pharmacy for the treatment, the lab to get the results of the labs..." **GP_3**

When the patient comes with a referral letter, the care actors pick information to contextualize their contribution and build the sequence of actions they need to put in place. However, overabundant information in the letters may scatter the attention of the actors from the important point, and consequently, they may focus on side issues. This problem occurs when care actors rely on their digital medical records

to generate parts of the referral letters. Thus, care actors have to decide what to inscribe on the referral letter.

“I keep telling myself that if I send my patients with a letter that is not precise enough, the specialists will not have all the information they need. Yet, if the letter is more elaborated, I know that they will not read everything. Sometimes this is what happens. We can see that they have not read everything because they answer vaguely, missing the main point” **GP_3**

To complement the referral letter, care actors further search for information about the patient. Communication may occur, in those cases, to be able to contextualize the information given in the documents. However, this communication's need and frequency depend on other circumstances; for example, a specialist receiving a letter to take care of type 2 diabetes does not necessarily reach out to the referral letter's sender. However, in a complex situation, communication is indispensable if the letter does not provide sufficient information. The verbatim below effectively explains this discrete communication's role in enhancing collaboration.

“For me, the collaboration is the exchange around a patient. Not always, not necessarily all the time, but punctually. For me, this is what collaboration is all about. It really is a punctual exchange.” **Podiatrist_2**

On the other hand, when care actors meet a patient they are used to taking care of, they simply rely on what they have already collected.

“To recall what I did with the patient the last time I saw them, I check the observations we had made before and the emails we received and integrated into our medical record.” **GP_3**

When creating summaries, care actors tend to select the information that helps them decide from their specialty's point of view. However, when they want to share these summaries with other care actors, they choose to provide only the information that might help the corresponding care actors:

"With diabetologists, systematically, on their prescriptions, they put the patient's history in few lines and below, they prescribe very well the specific sessions we need to provide to take care of diabetic patients. For example, a patient with a grade 2, which means that he has neuropathy or arteriopathy, will be entitled to four sessions in the year or one per quarter.... So, it is very, very well described" **Podiatrist_2**

Sharing summaries opens a door permitting care actors to access some part of their correspondents' medical records that they need to get an overview of their patients' trajectories. The verbatim below explains the importance of giving this access to parts of the medical record to enhance collaboration among the care circle functioning around the patient care pathway.

“The collaboration is to have access to some parts of the patient's medical record, the ability to have access to the different letters, summaries and prescriptions of people seeing the patient help us a lot to work together” **Endocrinologist_3**

However, the target of those summaries varies. The GPs, present the patients' pathways gatekeeper; thus, they tend to receive all the summaries generated by the different care actors. For other care actors, their function and role in the patient's trajectory define whether they receive summaries or not. In the integrated care pathway configuration, summaries tend to be shared with the medical actors (GP

and specialists) more than with the paramedic actors. However, occasionally including this category of actors is essential. In particular, sharing the produced summaries about the patient with the paramedic actors may help and direct their interventions as indicated by the lab doctor:

“We have a very light communication with the actors around us. Because in fact, we [the laboratory] send them the results. So, it goes in this direction; we send them the results for the patients they follow, and we send the results through faxes, ApiCrypt [secured medical mail service], or secured health messaging services. In return, we do not really have a return...but indeed, we want to have feedback even if this is punctual... generally, I call the attending physician. I try to check with him if he agrees to give me the diagnosis, especially in hematology. Knowing their patients' pathology makes it easier to analyze the results. When we examine cells, blood cells, it is easier for us. If we already know the diagnosis, if the diagnosis has already been made, we know better what to look for than if we have no information.” **Lab_doctor**

This vignette depicts the role of the laboratory doctor in a patient's trajectory as an information sender but never as a receiver. Yet, at the same time, it illustrates the need to be involved in the “sharing circle” as this may help them to refine their understanding and therefore get an overview of the patient's case supporting their analysis work.

Updating or complementing case overview

Care actors might decide to inform other care actors in the care pathway after a change in the patient's trajectory or the arrival of new information using different artifacts. For example, a patient starts a new therapeutic education procedure that requires other care actors to adapt their treatment.

“I do not necessarily write to the GP after each follow-up session. However, if the patient starts a new therapy program, such as outpatient insulin therapy, I send a letter to the GP. Therefore, I inform him that I see the patient in consultation within the framework of this program, and I inform him that the patient will be seen for a certain number of sessions. After that, I write to him again at the end of all the consultations to tell him how it went.” **Nurse_1**

Moreover, they communicate directly using synchronous channels to raise awareness about emergent problems and inform the concerned care actors about potential contingencies. For example, the podiatrist below explains how she reacts after identifying an emergent problem by calling the GP to ask him to take care of the patient or negotiate a solution.

“When there is an emergency, I inform the patient that he must seek treatment ... and I call his doctor. When I have the practitioner on the phone, the reaction is immediate; he may say [the correspondent] yes, I can take him [the patient] right away, or I cannot take care of him right away, tell him to come tomorrow so the next day. It is good to know immediately what to do. Or he may say, look, if you think it's too urgent, send him to the hospital.” **Podiatrist_2**

In addition, care actors exchange to report problems they have noticed during their consultation to other correspondents or book medical appointments for their patients. For example, the vignette below illustrates a case where a patient was sent

to the GP to start a treatment for a new issue she did not notice or treat before, following a discussion between two specialists seeing the same patient.

"I have a patient with high blood pressure levels who went to see a preventive doctor. It turns out that the patient told him [the preventive doctor] that I was his GP. The preventive doctor requested a tele-expertise consultation with a cardiologist from Nancy; they did the electro and found ventricular hypertrophy. Therefore, they decided it was important to send him back to his doctor, me, to start treating his blood pressure. And I looked [in the medical record] and noticed that I had not seen him for four years" **GP_1**

Thus, during emergencies, the care actors opt for phone calls to pass the information and ensure it is well-received by the right person.

"The easiest way is the phone. It is still the most direct and simple tool if you want to reach someone urgently." **Lab_doctor**

According to their current role in the care journey, care actors might also respond to requests asking to produce summaries to get information about the patient, as the physiotherapist put it.

"We, systematically, never make reports except when the GP asks. For example, last week, I had the doctor on the phone who asked me to take care of one of his patients because he has a minor intervertebral disorder. So, he gave me a very precise diagnosis of the problem, and he asked me for a check-up because the patient was in horrible pain. Therefore, I quickly booked an appointment with the patient. I had the GP on the phone at 11 a.m. at 2 p.m. I was taking charge of the patient, and I prepared an explanatory report, which I sent to the GP afterward." **physiotherapist**

— At the same time, Care actors converse not only to seek information but also to correlate their understanding of a patient's situation and identify potential gaps and information that they may lack, as put by the dietician:

"When I talk with psychologists, we sometimes have trouble crosschecking our information because we realize that patients, for example, when they want to have surgery [bariatric surgery], they have a precise idea in mind, and know what they need to avoid saying [so that the surgery is accepted]. So, they will give me some information and give the psychologist other information, or vice versa. However, we can easily identify that because we crosscheck the information we have from the patient, and we find out that they do not fit. So, checking the story we build about the patient with the others is very important" **Dietician_3**

The previous verbatim highlights the necessity to communicate to validate the understanding of each care actor of the patient situation. Besides correcting possible confusion, communication enables care actors to revise and share opinions and advice. Therefore, communication grants a space for each participant to contribute to each other's case overview.

Issues arising when trying to get a case overview

While care actors count on the patient and their documents to find their correspondents, identifying those actors becomes more complicated if they do not get those documents or if the patient cannot name their care ensemble. This problem is even more salient when the patient is admitted into a health facility. For

example, the following vignette describes the efforts of a GP working at a mental hospital to find the cardiologist treating her hospitalized patient.

"We have to go fishing... I mean, we call the attending physician. If the admission tells us, he sees a cardiologist, but no one knows which cardiologist, we call the cardiologists. If someone in Troyes follows him, we can find who that is after a few calls, but we do not get any answer if he is followed elsewhere. We try to see if someone in the local hospital follows him. Nevertheless, we do not have access to the email of those specialists (cardiologists working in the local hospital), so we are obliged to call in all the services and ask them one by one if they have any idea. It is very, very distressing." **GP_3**

Although identifying the correspondents facilitates seeking more information, getting information is not always guaranteed. For example, contacting care actors with private practice is more accessible than those employed in health organizations. In particular, care actors with private practice struggle to reach out to the care actors at the hospital, as they are only reachable via the hospital's switchboard operators. The vignette below illustrates the pharmacists' hardness when they call the hospital to get more information about the dosage inscribed in the prescriptions. While it is much easier to reach and exchange with the care actors working in the outpatient departments, contacting the actors working in the day hospital and the Emergency Department is more difficult due to work shifts.

"With the hospital, it is always complicated because you must find the right prescribing department. Often, the prescription comes from outpatient consultation, but sometimes it is the day hospital. In this case, we call the department written on the prescription. But the person who wrote the prescription is never there. Sometimes, it is not the right service. Sometimes you call, and you will be redirected to the switchboard operators [...] When the prescription comes from the emergency center, he [the prescription's writer] is a student, most of the time. Then we are almost sure he will not be at the hospital the next day of his nightshift; In this case, no one wants to take over the responsibility. It is always a little complicated; we struggle when there is a problem reaching someone in the hospital." **Pharmacist**

Therefore, the care actors with private practice rely on their networks of contact to reach care actors at the hospitals, hoping to find the person they are searching for to gather more information about the patient.

On the contrary, when communicating with actors with private practice, the availability of actors to take calls, give reviews, and respond to requests, varies. Several reasons may hinder communication, particularly the time constraint of medical activity. At the same time, the French reimbursement system does not include answering a request as a medical act. Consequently, taking different calls from colleagues may lead to a higher workload without being paid for it. Therefore, different care actors in the city avoid giving their mobile phone numbers and rely on filtering demands through their assistants. Besides the difficulty of getting more information about the patient, identifying and reaching out to the correspondents is more arduous during emergencies, especially during the evenings and the weekends. The lab doctor explains the struggle to pass urgent information during the weekends due to the unavailability of the doctors.

“The problem occurs on Saturdays and especially on Saturdays; the problem arises in communication. Generally, we have a lot of trouble joining doctors during the weekend. All the doctors have great difficulty reaching other doctors on Saturdays. Because on Saturday afternoon, there are very few doctors who work. Therefore, when there is a result [Lab result] on Saturday, it is really, really difficult to pass the information. Therefore, we try to reach either the patient or his nurse. However, when we cannot reach one of them, we call 15 [Urgent medical aid service], and we try to give them the information, hoping that they can do something and possibly go and see the patient in the worst case.” **Lab_doctor**

Moreover, when it comes to sharing the different types of documents, the care actors struggle to identify the receiving channels of their correspondents. For example, the vignette below describes how difficult it is for the endocrinologist to identify the right channel to send a document to each correspondent.

“The difficulty... I mean, the greatest difficulty today is to know the channels of reception of the person to whom we want to write. And that... I believe we will make considerable progress if we have just one simple unified tool. We really wish to write to them [correspondents] without spending time finding a way to do it. The secured messaging system is good, but there are several others. We do not necessarily know which one each person uses, and which one is operational. Several times, some doctors called me saying that they did not receive my mail... If we propose to communicate through a channel, sometimes I get the answer: I do not use it or look at it. In addition, there are secured messaging systems where there is no way to know if the mail has been received and read, a bit like normal mail. This is also difficult. On the other hand, we also know if the address is not good, so we have a returned message. But if we do not have a return, we stay on standby.” **Endocrinologist_4**

Indeed, the involvement of the different care actors, the medical and the social organizations, imply different reception channels depending on the organization's policy and budgets. Hence, the actors struggle to identify the proper channels to reach other care actors.

To summarize, to get an overview of a patient's situation, care actors couple documentation and communication practices, communication often aiming at contextualizing the information inscribed in the documents. These practices, conducted by each care actor, enhance the collective management of the patient situation as care actors have a better understanding of the case.

Discussion

Building a clear and solid understanding of a patient's situation is crucial in providing healthcare. Our results reveal that getting this overview in the context of integrated care is a complex process. Previous CSCW studies described this process as a collaborative activity in which a patient and several care actors contribute to developing their mutual understanding according to their knowledge (Jensen & Bossen, 2016; Mønsted et al., 2011). However, whereas previous work focused on the role of an integrated medical record in getting this overview and suggested complementing it with communication systems (Richter et al., 2016), our work confirms that both documentation and communication practices are essential in this

process. Furthermore, our study revealed that documents and communication are intertwined to get an up-to-date and efficient overview of the patients' cases. In this section, we discuss the role of each component in getting a case overview.

Documents as building blocks of cases' overviews

Our case study reveals that the documents forming the different parts of the medical record (core parts, procedural reports, and episodes' summaries) are central among care actors. Therefore, when trying to get an overview of a patient, care actors engage in a constant effort to reassemble information recorded within those parts. However, while we agree with (Chen & Pine, 2014) on the role of the patient as a messenger, and we have also noticed their role as an information source, our findings reveal that the contribution of the patient is not always efficient. The patient's incapacity to recall the actors and the steps they passed through during their health trajectory sends the actors on a blind search for contacts and information.

At the same time, by sharing documents, care actors deliver information about the landmark steps of the medical journey. For example, the referral document precedes a new care episode, and summaries give insight into the evolution of those episodes.

(Zhang et al., 2017) identified the challenge to keep the meaning of the traveling information undergoing the de-contextualization and then re-contextualization. In this case, our study identified the start of a conversation among actors following the reception of different documents to validate and consolidate the understating of the medical data.

Communication as a Cement to build a case overview

Communication and dialogue between team members are crucial to building a clinical overview (Jensen & Bossen, 2016). We argue that the conversation and dialogue come after analyzing the different parts of the medical record within integrated care.

The success of the integrated care pathways requires the collaboration between different geographically dispersed care actors with different specialties. While their work to manage the patient's care stretches over long periods, previous works debated the need to keep the ensemble of those care actors informed along the care journey to ensure coherent care within the integrated care (Stille et al., 2003). However, our study argued that the role of the care actors in the trajectory defines their need for information (to take the right action). For instance, although they are usually excluded from the systematic reception of medical records, paramedic actors still succeed in coordinating their work with other actors. At the same time, keeping the participants informed varies according to the type of information

exchanged, the urgency of the information, and the correspondent's role in the care pathway.

Moreover, our results reveal that communication is event-driven and generally problem-centered. Communication occurs to help care actors find information related to a problem facing, treat acute issues or validate the information collected. Thus, our results confirm (Abou Amsha & Lewkowicz, 2016) work as the organization of work within the integrated care is based on the emergent knots that appear following the need to treat issues.

At the same time, due to the complexity of patients' medical conditions treated through the care pathway, our case study shows that the discussions arising among the care actors are swirling around specific themes that trigger the exchanges. For example, we can see that the dietitian and the psychologist organize their conversation around the bariatrics element more than the other problems once the patient's enrollment in the surgical episode is started. Therefore, we suggest that enabling the visualization of the evolvement of those thematic-based discussions will enhance the construction of the overview.

Implication for design

To build and maintain their understanding of their patients' situations, the care actors need to be aware of the evolution of their patients' trajectories through the information located in documents that compose the patient records and the communications that emerge among those actors working in an integrated way. Therefore, we suggest that the new generation of medical records design should consider the coupling of documents and the discussions around them. In addition, we argue that enabling the care actors to navigate through this large amount of information will foster their collaboration and facilitate getting an overview. However, while the medical records allow the identification of the information's source and permit following the temporal evolution of the medical information, we have observed that care actor work practices are centered on the medical problems. Therefore, although previous problem-centered medical records failed to survive, we argue that designing a system based on emerging discussions around documents could be a new way of offering problem-based medical records. Moreover, considering the discussions will support the collective sense-making of data around problems mandatory for the care actors to deliver care (Paul & Reddy, 2010).

Moreover, while the previous work on medical visualization aimed to establish visualization systems that support individual diagnosis and decision-making, we argue for the potential of visualization techniques to get the chronological evolution of documents and discussions and, therefore, to picture the patient's trajectory. Being able to follow and identify the relevant and recent documents and conversations around a patient would help the actors contextualize the evolution of the patient's situation and then get a clear overview of the case.

We suggest that the visualization of medical documents should then combine three dimensions: the source, the time, and the problem. The time and source dimensions would allow the care actors to follow the temporal evolution of the patients' cases. While the problem dimension will allow navigating the case through pined exchanged documents and the essential discussions arising around those documents.

Conclusion

In this paper, we investigated how care actors build, update, and maintain an overview of a patient's situation that they rely on to deliver care in the context of integrated care. Our case study confirms the importance of the medical information shared through documents and the direct communication among actors to get this overview. It also offers more details on how an overview can be achieved; while the documents shared among the care actors involved in the integrated care pathways are the cornerstone, discrete communication is essential to consolidating and sharpening the understanding. Based on these findings, we suggest designing problem-centered medical records with a visualization that helps actors get an overview of cases. Inspired by our observed practices, we suggest that this visualization offers three dimensions: time, source of information, and problem.

Acknowledgment

This work is supported by a grant from the Grand East region and the Aube department.

We would like to thank the staff of "Respo'Aube", for their support.

References

- Abou amsha, K., Bossen, C., Grönvall, E., & Lewkowicz, M. (2020). Computer-Supported Knotworking: Design Guidelines based on Two Case Studies from the Healthcare Domain in Europe. *Proceedings of the ACM on Human-Computer Interaction*, 5, 29.
- Abou Amsha, K., & Lewkowicz, M. (2016). Shifting patterns in home care work: supporting collaboration among self-employed care actors. *COOP 2016: Proceedings of the 12th International Conference on the Design of Cooperative Systems, 23-27 May 2016, Trento, Italy*, 139–154.
- Abou Amsha, K., & Lewkowicz, M. (2018). Chapter 3 - Supporting Collaboration to Preserve the Quality of Life of Patients at Home—A Design Case Study. In M. S. Ackerman, S. P. Goggins, T. Herrmann, M. Prilla, & C. Sary (Eds.), *Designing Healthcare That Works* (pp. 39–57). Academic Press. <https://doi.org/10.1016/B978-0-12-812583-0.00003-1>
- Abraham, J., & Reddy, M. C. (2013). Re-coordinating activities: an investigation of articulation work in patient transfers. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work - CSCW '13*, 67. <https://doi.org/10.1145/2441776.2441787>

- Amir, O., Grosz, B. J., Gajos, K. Z., Swenson, S. M., & Sanders, L. M. (2015). From Care Plans to Care Coordination: Opportunities for Computer Support of Teamwork in Complex Healthcare. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, 1419–1428. <https://doi.org/10.1145/2702123.2702320>
- Berg, M. (1996). Practices of reading and writing: the constitutive role of the patient record in medical work. *Sociology of Health & Illness*, 18(4), 499–524.
- Berg, M. (1999). Accumulating and coordinating: occasions for information technologies in medical work. *Computer Supported Cooperative Work (CSCW)*, 8(4), 373–401.
- Bossen, C., & Jensen, L. G. (2014). How physicians' achieve overview? a case-based study in a hospital ward. *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 257–268.
- Chao, C.-A. (2016). The impact of electronic health records on collaborative work routines: A narrative network analysis. *International Journal of Medical Informatics*, 94, 100–111. <https://doi.org/10.1016/j.ijmedinf.2016.06.019>
- Chen, Y., & Pine, K. (2014). When Medical Expertise Meets Record Expertise: The Practices of Patient Accessible Medical Records in China. In C. Rossitto, L. Ciolfi, D. Martin, & B. Conein (Eds.), *COOP 2014 - Proceedings of the 11th International Conference on the Design of Cooperative Systems, 27-30 May 2014, Nice (France)* (pp. 409–426). Springer International Publishing. https://doi.org/10.1007/978-3-319-06498-7_25
- Clynch, N., & Kellett, J. (2015). Medical documentation: part of the solution, or part of the problem? A narrative review of the literature on the time spent on and value of medical documentation. *International Journal of Medical Informatics*, 84(4), 221–228.
- Coiera, E. (2000). When Conversation Is Better Than Computation. *Journal of the American Medical Informatics Association*, 7(3), 277–286. <https://doi.org/10.1136/jamia.2000.0070277>
- Cresswell, K. M., Robertson, A., & Sheikh, A. (2012). Lessons learned from England's national electronic health record implementation: implications for the international community. *Proceedings of the 2nd ACM SIGHIT Symposium on International Health Informatics - IHI '12*, 685. <https://doi.org/10.1145/2110363.2110441>
- Cresswell, K. M., Worth, A., & Sheikh, A. (2012). Integration of a nationally procured electronic health record system into user work practices. *BMC Medical Informatics and Decision Making*, 12.
- Crocker, A., Trede, F., & Higgs, J. (2012). Collaboration: What is it like?—Phenomenological interpretation of the experience of collaborating within rehabilitation teams. *Journal of Interprofessional Care*, 26(1), 13–20.
- Embi, P. J., Weir, C., Efthimiadis, E. N., Thielke, S. M., Hedeem, A. N., & Hammond, K. W. (2013). Computerized provider documentation: findings and implications of a multisite study of clinicians and administrators. *Journal of the American Medical Informatics Association*, 20(4), 718–726.
- Feldman, M. S., Sköldböck, K., Brown, R. N., & Horner, D. (2004). Making sense of stories: A rhetorical approach to narrative analysis. *Journal of Public Administration Research and Theory*, 14(2), 147–170.
- Fitzpatrick, G., & Ellingsen, G. (2013). A Review of 25 Years of CSCW Research in Healthcare: Contributions, Challenges and Future Agendas. *Computer Supported Cooperative Work (2013)*, 57. <https://doi.org/DOI.10.1007/s10606-012-9168-0>
- Fukuda, H., Shimizu, S., & Ishizaki, T. (2015). Has the Reform of the Japanese Healthcare Provision System Improved the Value in Healthcare? A Cost-Consequence Analysis of Organized Care for Hip Fracture Patients. *PLOS ONE*, 10(7), e0133694. <https://doi.org/10.1371/journal.pone.0133694>
- Gulmans, J., Vollenbroek-Hutten, M. M. R., Van Gemert-Pijnen, J. E. W. C., & Van Harten, W. H. (2009). Evaluating patient care communication in integrated care settings: application of a mixed method approach in cerebral palsy programs. *International Journal for Quality in Health Care*, 21(1), 58–65. <https://doi.org/10.1093/intqhc/mzn053>

- Gurses, A. P., & Xiao, Y. (2006). A Systematic Review of the Literature on Multidisciplinary Rounds to Design Information Technology. *Journal of the American Medical Informatics Association*, 13(3), 267–276. <https://doi.org/10.1197/jamia.M1992>
- Hayrinen, K., Saranto, K., & Nykanen, P. (2008). Definition, structure, content, use and impacts of electronic health records: A review of the research literature. *International Journal of Medical Informatics*, 77(5), 291–304. <https://doi.org/10.1016/j.ijmedinf.2007.09.001>
- Jensen, L. G., & Bossen, C. (2016). Creating overview via paper-based and electronic patient records. *Scandinavian Journal of Information Systems*, 28(2), 26.
- Kodner, D. L., & Spreeuwenberg, C. (2002). Integrated care: meaning, logic, applications, and implications—a discussion paper. *International Journal of Integrated Care*, 2.
- Lutters, W. G., & Ackerman, M. S. (2002). Achieving safety: A field study of boundary objects in aircraft technical support. *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work*, 266–275.
- Marref, R., Abou Amsha, K., & Lewkowicz, M. (2022). Practice-Centered Approach to Design Cooperative Healthcare Information Systems: Data, Architectural and Organizational Challenges. *COMPUTING AND INFORMATICS*, 41(1), 191–212.
- McDonald, C. J., & Barnett, G. O. (1990). Medical-record systems. In *Medical informatics: computer applications in health care* (pp. 181–218).
- Mønsted, T. (2015). Keeping Distributed Care Together: Medical Summaries Reconsidered. *ECSCW 2015: Proceedings of the 14th European Conference on Computer Supported Cooperative Work, 19-23 September 2015, Oslo, Norway*, 143–161. https://doi.org/10.1007/978-3-319-20499-4_8
- Mønsted, T., Reddy, M. C., & Bansler, J. P. (2011). The Use of Narratives in Medical Work: A Field Study of Physician-Patient Consultations. In S. Bødker, N. O. Bouvin, V. Wulf, L. Ciolfi, & W. Lutters (Eds.), *ECSCW 2011: Proceedings of the 12th European Conference on Computer Supported Cooperative Work, 24-28 September 2011, Aarhus Denmark* (pp. 81–100). Springer. https://doi.org/10.1007/978-0-85729-913-0_5
- Paul, S. A., & Reddy, M. C. (2010). Understanding together: sense-making in collaborative information seeking. *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, 321–330. <https://doi.org/10.1145/1718918.1718976>
- Pescosolido, B. A. (2013). Patient Trajectories. In W. C. Cockerham, R. Dingwall, & S. Quah (Eds.), *The Wiley Blackwell Encyclopedia of Health, Illness, Behavior, and Society* (pp. 1770–1777). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118410868.wbehibs282>
- Plaisant, C., Milash, B., Rose, A., Widoff, S., & Shneiderman, B. (1996). LifeLines: visualizing personal histories. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 221–227.
- Powsner, S. M., & Tufte, E. R. (1994). Graphical summary of patient status. *Lancet*, 344(8919), 386–389.
- Reddy, M. C., Dourish, P., & Pratt, W. (2001). *Coordinating heterogeneous work: Information and representation in medical care*. Kluwer Academic Publishers, Dordrecht, The Netherlands. <https://doi.org/10.1007/0-306-48019-13>
- Richter, P., Gand, K., & Esswein, W. (2016). Lifetime Electronic Health Records as an Enabler for Integrated Care. *2016 49th Hawaii International Conference on System Sciences (HICSS)*, 3199–3207. <https://doi.org/10.1109/HICSS.2016.401>
- Rind, A., Wang, T. D., Aigner, W., Miksch, S., Wongsuphasawat, K., Plaisant, C., & Shneiderman, B. (2013). Interactive Information Visualization to Explore and Query Electronic Health Records. *Foundations and Trends® in Human-Computer Interaction*, 5(3), 207–298. <https://doi.org/10.1561/11000000039>
- Stille, C. J., Korobov, N., & Primack, W. A. (2003). Generalist-subspecialist communication about children with chronic conditions: an analysis of physician focus groups. *Ambulatory Pediatrics*, 3(3), 147–153.
- Tange, H. (1996). How to approach the structuring of the medical record? Towards a model for flexible access to free text medical data. *International Journal of Bio-Medical Computing*, 42(1–2), 27–34.

- Tange, H. J., Hasman, A., de Vries Robbé, P. F., & Schouten, H. C. (1997). Medical narratives in electronic medical records. *International Journal of Medical Informatics*, 46(1), 7–29. [https://doi.org/10.1016/S1386-5056\(97\)00048-8](https://doi.org/10.1016/S1386-5056(97)00048-8)
- Vanhaecht, K., Sermeus, W., Peers, J., Deneckere, S., Lodewijckx, C., Leigheb, F., & Panella, M. (2010). The European Quality of Care Pathway (EQCP) Study: history, project management and approach. *International Journal of Care Pathways*, 14(2), 52–56.
- Weed, L. L. (1968). Medical Records that guide and teach. *N. Engl. J. Med*, 652–657.
- Zhang, Z., Sarcevic, A., & Bossen, C. (2017). Constructing common information spaces across distributed emergency medical teams. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 934–947.

Schubert, Petra and Williams, Susan P. (2022): Enterprise Collaboration Platform Configurations: an Empirical Study. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies – Notes, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_n01

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_n01

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

Enterprise Collaboration Platform Configurations: an Empirical Study

Petra Schubert and Susan P. Williams

University of Koblenz

schubert@uni-koblenz.de, williams@uni-koblenz.de

Abstract. Collaboration and communication technologies are essential for the support of cooperative work in organisations. Unlike the situation with ERP Systems, there is no single integrated Enterprise Collaboration System that provides systematic and comprehensive support for all the different forms of collaborative activities. As a consequence, organisations must combine multiple tools, applications and systems to build their Enterprise Collaboration Platform. In this paper, we present the findings of a focused empirical study that examines the complex collaborative technology landscape in user organisations in order to characterise and understand the evolving portfolios of collaboration software that have been implemented. Based on a literature review combined with an analysis of existing commercial software products, we developed a classification scheme for Areas of Collaborative Work (ArCoW), which is then used to structure an online questionnaire. The analysis of data from 23 responding user companies revealed three typical “configurations” of Enterprise Collaboration Platforms: *concentration*, where the platform is highly focused on a core ECS/suite with only a few additional collaboration software tools, *diversity* that also builds around a core ECS/suite but extends this with a wide range of additional tools and *dual core* characterised by two ECS/Suites with few additions.

Motivation

The use of collaboration technologies to support and coordinate distributed collaborative work is increasing in both scale and scope. This has been especially noticeable in the past two years as organisations have responded to the short-term challenges presented by the work from home mandates triggered by the COVID-19 pandemic (McKinsey Global Institute, 2021; World Economic Forum, 2020) as well as the longer-term trend for organisations to offer employees more flexible work options to support hybrid work (Gratton, 2021). This is resulting in a wider range of working arrangements and an increased requirement to provide collaboration technologies to support distributed work teams.

As a response to this increase in the scale of distributed collaborative working there has been a concomitant increase in the scope and functionality of technologies to support collaborative work (Gartner, 2021). As the degree of hybrid working increases, the complexity of supporting collaborative work has also increased, requiring IT departments to place greater attention on the selection and provision of “collaboration software” to provide a stable technology environment to support a wider range of collaborative work situations.

However, collaboration software designed to support *ad hoc collaborative work activities* is inherently different from software such as Enterprise Resource Planning (ERP) systems that, by design, support *repetitive and highly structured business processes*. Whilst the majority of today’s ERP Systems have evolved into highly integrated systems with functional modules that store data in a central database, the commercial solutions for collaboration software are highly specialised, focusing on supporting selected areas of joint work. Only a small number of collaboration software products (e.g. HCL Connections, Atlassian Confluence, Jive) contain multiple functional modules and would qualify as *integrated* Enterprise Collaboration Systems.

The fact that collaboration software is highly specialised means that organisations often need to combine many different software products to support the diverse requirements for joint work. This has led to a (somewhat uncontrolled) growth of available products in use in organisations (Schubert & Williams, 2022) and as a consequence, many companies have implemented a heterogeneous range of tools (from different vendors) with overlapping (redundant) functionality. These portfolios of collaboration software include lightweight tools for specific tasks such as file sharing or simple message exchange to more complex Enterprise Collaboration Systems (ECS). Taken together the tools provide an “Enterprise Collaboration *Platform*” (ECP) that comprises the full range of collaboration software tools and applications available to the registered users of an organisation. Frequently, the selection of these products arises bottom-up in a piecemeal fashion as different tools are requested by single departments and is not conduct-

ed top-down, centrally coordinated by the IT department (Riemer et al., 2012). Examining the portfolio of software products across the entire organisation reveals a diverse range of tools in use, provided by multiple vendors and supporting differing and often redundant functionality. The aim of our research is to examine whether typical “configurations” for Enterprise Collaboration Platforms are emerging to support distributed collaborative work. That is, our goal is to identify whether there are frequently occurring combinations of collaboration software products that, in their combination, provide identifiable collaboration platform types.

In this research note, we report on the findings of an empirical study of existing Enterprise Collaboration Platforms in 23 medium- to large-sized organisations. In order to investigate the design of Enterprise Collaboration Platforms, we develop a generic template for the **Areas of Collaborative Work (ArCoW)** and use this to examine similarities in existing collaboration tool portfolios reported by the participating user organisations and to identify *typical* configurations. The study builds on and extends our previous work on the use of social software in organisations (Schubert & Williams, 2022; Williams & Schubert, 2018) and is part of IndustryConnect, a long-term university-industry research programme exploring collaboration technologies and the design of the digital workplace in 35 leading organisations in the DACH area (Williams & Schubert, 2017). The research programme has been following the emergence and shaping of the digital workplace in these organisations since 2013 through the development of longitudinal case studies, research workshops, interviews and trace ethnographies of work and work practices.

The paper is organised as follows. In the next section, we introduce the research design and explain the development of the survey instrument and the collection of research data. We then present our analysis, and emerging platform configurations are identified. Finally, we conclude with a discussion of our findings and present imperatives for future research.

Research Design and Survey Instrument

This section describes the research design and the development of the survey instrument. As described above, our aim is to examine the emergence of enterprise collaboration platform types by identifying and analysing the portfolios of collaboration tools in use in organisations.

A mixed method research approach using an exploratory-sequential design was applied (Creswell & Plano Clark, 2018), which combines literature and software analysis with empirical methods for the collection and analysis of survey data. The *survey instrument* was developed based on a literature review followed by a thorough examination of the functionality of existing standard software solutions

(Schubert & Williams, 2022). The *survey data* was collected from user organisations through an online questionnaire.

Figure 1 shows the research steps and the outcomes of each step. In the *first step*, the findings from an analysis of *dimensions of collaborative work* in the CSCW research literature and the examination of *functionality of existing collaboration software* were merged to develop a classification of **Areas of Collaborative Work (ArCoW)**. In the *second step*, the ArCoW structure was used to design the online questionnaire and in the *third step*, the questionnaire was implemented using LimeSurvey (an open source online statistical survey Web application) and used for data collection.

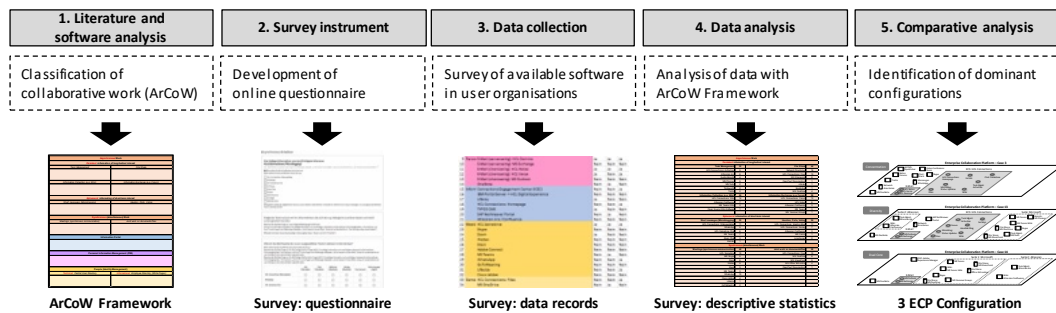


Figure 1. Research steps and results

The collected data records containing the portfolio of software in use in the responding user organisations were transferred into an Excel file for analysis. In the *fourth step*, the ArCoW Framework was used to derive descriptive statistics about the type and number of tools in use in the responding organisations. The *fifth and final step* was a comparative analysis of the data records which reveals similarities in the data that could be described as *typical configurations* of an Enterprise Collaboration Platform.

The following sections describe the research carried out in each of the steps in more detail.

Literature and Software Analysis: Areas of Collaborative Work

Investigating the infrastructure for *all areas* of collaborative work is a complex task and requires a classification scheme that is consistent with analytical frameworks developed in the academic literature as well as with “the language of practice”. Moreover, the classification must be applicable to examining the functionality bundles of contemporary commercial collaboration software. For this, we purposefully combined findings from a literature review with a functional examination of existing software products.

Literature review

For the development of our classification scheme, we examined existing approaches in the literature. We were particularly interested in meta classification schemes that provide a comprehensive classification of the entire field of collaborative work. Examples are the study of *categories of tools* by Bafoutsou and Mentzas (2002), Riemer's *catalogue of classification criteria* (Riemer, 2007), the *8C Framework* by Williams (Williams, 2010) and Schubert's *classification of software components* (Schubert, 2018). All of the identified classification approaches contain similar dimensions, the most prominent being *synchronicity* (synchronous/asynchronous) and *place* (Ellis et al., 1991), *permanency* of the information (ephemeral/long-term) (Schubert, 2018), type of *group process* (communication, cooperation, coordination) (Williams, 2010), *content type* (text, image, video, audio) and number of *communication partners* (1/many) (Ellis et al., 1991). The type of activity is often grouped into three (or four) "Cs": Communication, Cooperation, Coordination (and Content). In addition, studies have shown that collaboration software can be classified into categories, which are dependent on the *type of work* which they support. Riemer (2007) suggested the five categories *Integrated Systems*, *Everyday Systems*, *Meeting Systems*, *Coordination Systems* and *Specialised Tools* as a result of a cluster analysis on a sample of 94 collaboration software products using some of the attributes listed above. Bafoutsou and Mentzas (2002) found similar categories in their functional review of collaborative systems. Their main groups are *real-time conferencing*, *non-real time conferencing*, *file and document sharing*, *electronic workspace* and *electronic meetings systems*.

Software analysis

We used the *Collaborative Technologies Evaluation Tool* by Schubert and Williams (2011) for our analysis of leading commercial software products. The evaluation tool is based on the 8C Model (Williams, 2010) and contains a list of 42 functional criteria which are grouped by the four inner Cs (see Annex).

The preceding analysis of the literature showed that there is no accepted "standard" classification scheme for all areas of collaborative activity. The following analysis of the software identified that there is no *single integrated* Enterprise Collaboration System that covers all aspects of collaboration. Instead, the market for collaboration software is heterogeneous, comprising a multitude of commercial collaboration tools with both overlapping and disjoint functionality.

Based on this analysis, we developed and used a classification scheme that is relatively straightforward in its reflection of daily work practices and allows organisations to clearly define requirements according to user activity. To map the software functionality to this scheme, we decomposed software products on a modular level, assigning (some of) them to multiple areas. This turned out to be a feasible approach and it also helped to clearly identify overlapping and redundant

functionality. The resulting classification contains 8 *functional categories* (Figure 2) which represent the areas of collaborative work (ArCoW).

On the highest level, we distinguish between synchronous and asynchronous activity. *Synchronous* activity is further sub-divided into communication-oriented *meetings* and cooperation-oriented work on *documents/files*. *Asynchronous* activity has two additional sub-areas according to the *permanency* of the information: ephemeral (of short-term interest) and persistent (of long-term interest). The *ephemeral* section contains communication-oriented short messages (microblogging) as well as the coordinative features *ideation, polls and voting*, content which is normally only relevant for a short amount of time. The *persistent* section has four sub-areas: the coordination-oriented *task management* and the three content-oriented *file sharing, information collection* (documentation) and *information exchange* (question-response).

Asynchronous Work	
Persistent information of longitudinal interest	
Task Management	File Sharing
Information Collection	Information Exchange
Ephemeral information of short-term interest	
Short messages (Microblogging)	Ideation, Polls, Voting
Synchronous (simultaneous) Work	
Meetings	Joint work on documents/files

Figure 2. ArCoW Framework (Areas of Collaborative Work)

All dimensions that were identified in the literature review are contained in this classification but they are not all equally important/visible. Our main dimensions are *synchronicity* and *permanency* of the information. The sub-areas reflect different types of *group processes* (communication, cooperation, coordination, content creation). *Content type* (text, image, video, audio) and number of *communication partners* are implicitly embedded in the functionality provided by the software. We excluded the dimension of *place* (Ellis et al., 1991) in our preliminary analysis of distributed remote and hybrid work, where nothing is (solely) co-located. This is not to say that place is not important, however, for the purposes of this preliminary study where the objective is to analyse the constellations of software tools in use and emerging technology platforms being formed, the analysis of the place where someone is working from is less important than the mode of working, i.e. synchronous/asynchronous and the requirements (or not) for persistence of information. Place (and physical distribution of actors), as well as other dimensions such as those identified by Lee and Paine (2015) such as scale, scope and nature of work play an important part in our subsequent analysis of the collaborative work and work practices.

Introducing the platform view

The aim of our study was to investigate how organisations build their Enterprise Collaboration Platform – the collective portfolio of collaboration software to provide the technology environment to support enterprise-wide collaborative work. As demonstrated in our software analysis, commercial collaboration software can be assigned to specific areas of collaborative work (discussed above). When examining the infrastructure of a whole organisation (*the platform level*), we need to also consider the *basic technology stack* that is used to form the foundation of the platform. The foundation includes the *user management*, *personal information management* as well as the *organisational “homepage”* (usually a *portal or intranet*) that provides a structured and uniform access to the information resources of the organisation.

The *technical* aspect of the *Central User Directory* is covered by solutions for identity and access management (IAM) that comprise services for authentication, authorisation, user management as well as a central user repository. IAM provides the possibility for single sign-on for multiple software applications from desktops as well as mobile devices (Gartner, 2022b). The *informational employee directory* addresses the need to know the background of a person (area of expertise) in order to identify experts or appraise their contributions (name of person, room, contact info, expertise, role, ...). In recent years, this functionality has been added in the form of “Enterprise Social Networks” (ESN) (Wehner et al., 2016) or “Social Intranets” (Williams & Schubert, 2018). These terms were coined when “Enterprise Social Software” was introduced into organisations as a direct result of the success of the public Social Media (Leonardi et al., 2013). These two software types have similar functionality but differ in their primary objective. Both provide “social features” (social profile, link, follow, like, tag, post, comment, ...) where the *ESN* has a focus on people with the aim of establishing links between them to build an organisational network structure and the *Social Intranet* has a focus on content to share and increase awareness about information.

Personal Information Management (PIM) is the term used for the realm of information creation and organisation of each employee. PIM software enables individuals to create digital content (texts, slides, worksheets, graphics, databases, charts, videos, music, ...). The majority of documents are first created within individual desktop environments and only later “become social” when they are uploaded or copied into collaboration software. Content that is natively created in a collaboration software is “born social” (Hausmann & Williams, 2016).

Information Portals are typically implemented by means of *Digital eXperience Platforms (DXP)* (Gartner, 2022a) or *Content Service Platforms (CSP)* (Gartner, 2022c).

These basic platform components are the necessary foundation for the building of an Enterprise Collaboration Platform. Collaboration software for the different areas of collaborative work (ArCoW) is then added to the ECP according to the

requirements of the user organisation. Figure 3 shows an example of an ECP with its portfolio of software products. As mentioned earlier, software products containing multiple functional components can appear in multiple areas. For example, NextCloud is a tool for (asynchronous) *file sharing* but also allows users to *synchronously work on files*. HCL Connections is an integrated Enterprise Collaboration System with multiple functional modules (wiki, blog, forum, etc.) that supports almost all areas of asynchronous work. For our analysis, we decomposed such systems and suites into their separate modules. This explains why the same symbol can occur multiple times with different labels.


























Asynchronous Work	
Persistent information of longitudinal interest	
Task Management	File Sharing
  HCL Connections: Activities HCL Kudos Boards	   Network Directories HCL Connections: Files NextCloud
Information Collection	Information Exchange
 HCL Connections: Wiki	 HCL Connections: Forum
Ephemeral information of short-term interest	
Short messages (Microblogging)	Ideation, Polls, Voting
   HCL Connections: Status updates HCL Sametime MS Skype	   HCL Connections: Survey HCL Connections: Ideation Blog LimeSurvey
Synchronous (simultaneous) Work	
Meetings (synchronous communication)	Joint work on documents/files
 Zoom	   HCL Connections: Files NextCloud Office 365 online
Information Portal	
Intranet	Portal
 HCL Connections: Homepage	
Personal Information Management (PIM)	
E-Mail	Productivity Tools
   HCL Notes (E-Mail client-side) HCL Domino (E-Mail server-side) Sogo (E-Mail server-side)	 MS Office
People (Identity Management)	
Technical: Central User Directory	Informational: Employee Directory (White Pages)
  HCL Domino LDAP LDAP	 HCL Connections: Social Profiles

Figure 3. Case example for a platform configuration (ArCoW and basic platform components)

Data Collection and Analysis

In the survey, we investigated the current portfolios of collaboration software that user companies have implemented to build their ECP. Data was gathered from a sample based on responses from 23 user companies that, taken together, have a total of more than 730,000 employees. The user organisations are all members of the research initiative IndustryConnect (Williams & Schubert, 2017) and are interested in supporting collaboration research and the design of the Digital Workplace. They are mostly large organisations located in the DACH area, and they represent different industry sectors (e.g. manufacturing, engineering, services).

Asynchronous Work			
Persistent information of longitudinal interest			
Task Management	44	File Sharing	52
HCL Connections: Activities	12	Network directories	19
Jira	12	HCL Connections: Files	15
MS Planner	7	MS Sharepoint	10
Kudos Boards	5	MS OneDrive	5
MS Todo	4	OwnCloud/Nextcloud	1
Trello	2	OpenText Documentum	1
Taskworld	2	nScale	1
Information Collection	31	Information Exchange	29
HCL Connections: Wiki	16	HCL Connections: Forum	15
Atlassian Confluence	10	HCL Notes	6
MS Teams Wiki	4	MS Teams	4
Open Source Wiki-Software	1	Atlassian Confluence	3
		MS Yammer Groups	1
Ephemeral information of short-term interest			
Short messages (Microblogging)	38	Ideation, Polls, Voting	28
HCL Connections: Statusupdates	13	HCL Connections: Ideation Blog	12
MS Teams	10	HCL Connections: Survey	11
HCL Sametime Chat	6	MS Forms	3
WhatsApp	5	LimeSurvey	2
Zoom Chat	2		
Kaizala	2		
Synchronous (simultaneous) Work			
Meetings (synchronous communication)	64	Joint work on documents/files	33
MS Teams	14	HCL Connections: Files	12
HCL Sametime	10	MS OneDrive	10
Skype	7	SharePoint	6
Zoom	7	Office 365 online	4
Webex	7	OpenText Documentum	1
WhatsApp	5		
GoToMeeting	5		
Slack	4		
Lifesize	3		
Cisco Jabber	2		

Figure 4. Software products in use by the surveyed user organisations (n=23)

The questionnaire respondents are the people in these organisations who are responsible for identifying and supporting the user requirements regarding collaboration software. In the questionnaire the respondents were asked to identify the products that are contained in their organisation's portfolio from a list of software products for the different areas of Enterprise Collaboration. It was also possible

for respondents to add further software tools in a free text field if these were not included on the original list.

The data from all the respondents was consolidated and an inventory of all the tools reported by organisations was created and then analysed. Figure 4 contains an overview of the functional modules that were selected (or added) by the respondents. The numbers in Figure 4 show that the organisations involved in the study have implemented a substantial number of *different* commercial software products to build their Enterprise Collaboration Platform (ECP). 319 different software modules (for ArCoW) were identified in the sample of 23 organisations, which represents an *average of 14 functional modules per organisation*. The numbers reveal the high degree of redundancy in the available functionality. For example, the 23 organisations have implemented an average of 2,78 applications per organisation for synchronous meeting support (64 in total).

Technology Landscape of an Enterprise Collaboration Platform (ECP)

For the visualisation of the Enterprise Collaboration Platforms of the study participants, we developed a template for visualising a “platform image” that shows the areas of the classification. We then used this template for the visualisation of the 23 company data sets. Figure 5 shows an example platform visualisation for one of the cases. As explained earlier, the software used to provide the platform foundation is displayed on the bottom of the graphic. Most of the respondents use Microsoft Active Directory or some form of LDAP for their technical user directory and the majority have implemented an informational user directory, either in the form of an ESN or Social Intranet.

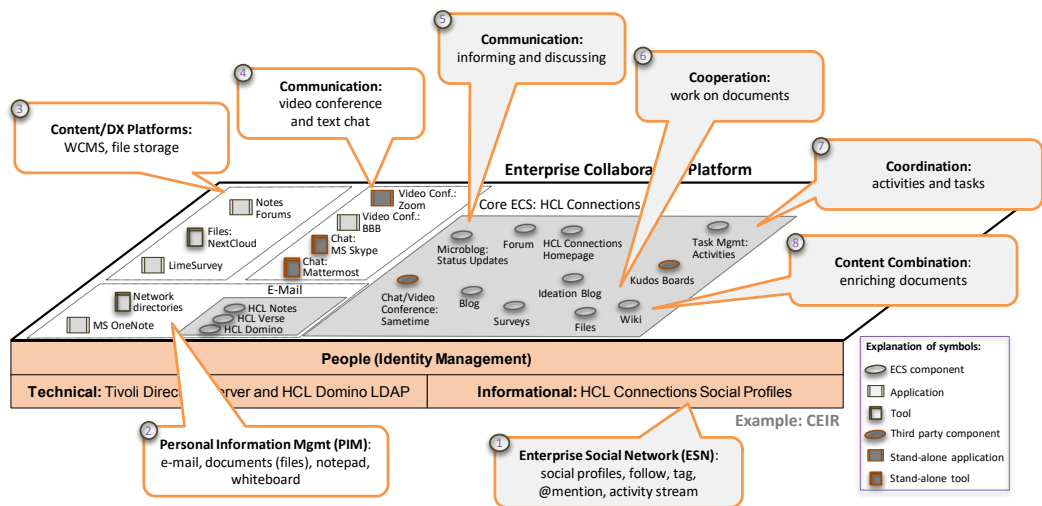


Figure 5. Case example of an Enterprise Collaboration Platform: modules grouped by ArCoW

The software modules are clustered in groups according to their category. The symbols indicate if the software is a tool, an application or a component of a sys-

tem or suite (see legend). The dark colour indicates that the software is “stand-alone” and uses its own technical user directory (no single sign-on). The tools for Personal Information Management are shown on the bottom left. We only included selected tools such as E-Mail and personal notebooks that have collaborative functionality. Tools for portals and intranets are placed on the top left. As already mentioned, the study showed that the area of (near) synchronous communication contains a wide variety of specialised solutions for video conferencing and text chat which is why we allocated it to a specific section.

The rest of the platform contains the modules that support the *core ArCoW modules* and this was our candidate section for the identification of typical configurations.

The software products used to build the ArCoW core differ in the *range of functionality* and the *degree of integration* between the modules. The following four types could be identified: *Enterprise Collaboration System/ECS (1)* are a purposefully developed selection of applications/tools that are fully integrated and provided to the user in a workspace under a uniform interface (e.g. HCL Connections). A *Collaboration Suite (2)* is a bundle of applications/tools (often under a joint license) that can be used independently. They provide a certain degree of technical integration because they have been designed to work together (e.g. the collaboration suites by Google, Atlassian and Microsoft). An *application (3)* is a standalone software product with multiple collaboration features (e.g. TeamViewer with screen sharing, video conferencing and file transfer). A *tool (4)* is a lightweight desktop or mobile software/plugin/functionality with a central focus on one/few features (e.g. chat in WhatsApp).

The analysis of the survey data revealed three dominant ECP configurations, defined as *Concentration*, *Diversity* and *Dual Core*. A small number of companies were still in the early stages of defining their collaboration portfolio and were classified into a group named “*Forming*”.

Dominant Platform Configurations

The ECP framework was used to synthesise the survey data about the types and the range of collaboration technology in use in the responding organisations (the “cases”) and the collaboration technology landscape for each case was visualised. The results were then examined to gain insights into different collaboration platform configurations.

The three dominant configurations are:

- (1) **Concentration** (core ECS/suite with a **few** additions)
- (2) **Diversity** (core ECS/suite with a **broad range** of additions)
- (3) **Dual Core** (**two** ECS/Suites with a few additions)

Figure 6 shows a visual comparison of the assemblages of tools comprising the emerging enterprise collaboration platforms for three selected case companies.

The visualisations show the different enterprise collaboration platform design strategies that the organisations are following.

Case Company 3 is following the **Concentration approach**. The company is using all functional modules of the integrated ECS. Only where functionality is missing or insufficiently provided, are they making use of complementary tools (e.g. in this case the company is doing this to support synchronous communication through video conferencing and specialised task management through Trello).

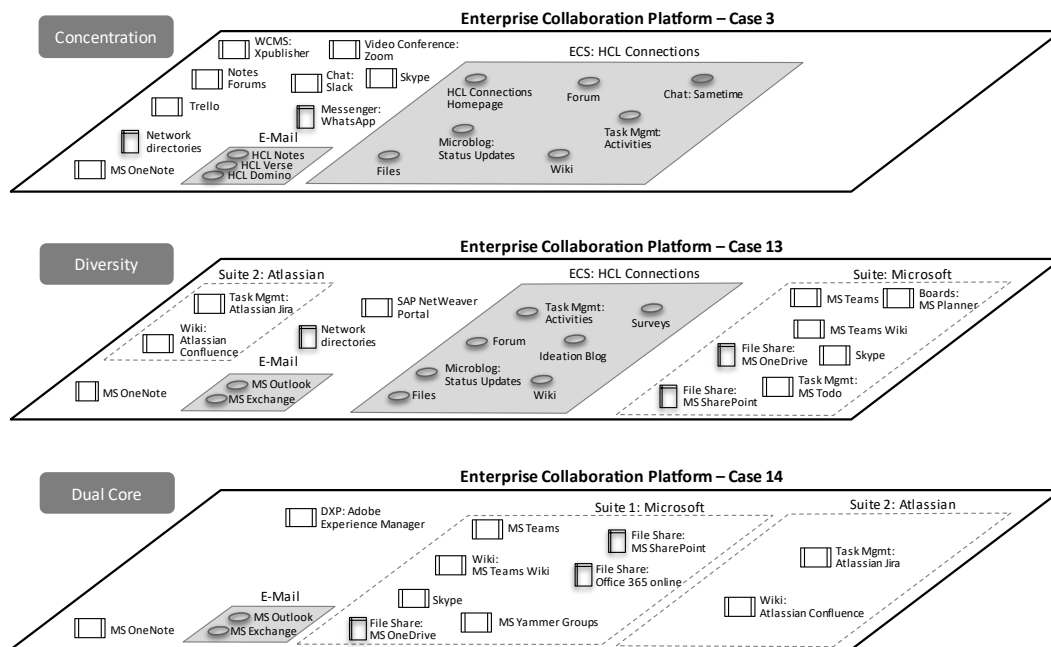


Figure 6. Three dominant ECP configurations: Concentration, Diversity and Dual Core

Case Company 13 is following the **Diversity approach**. The company is only using selected modules in the core ECS and is complementing the functionality of the platform with applications from two other collaboration suites. Whilst this configuration provides the users with a broad range of options for collaboration support it also creates the problem of redundancy (e.g. in this case there are three different software products supporting wikis).

Case Company 14 is following the **Dual Core approach**. The company is using two suites (Microsoft and Atlassian) to provide the necessary ArCoW functionality.

There are many commonalities in our sample in the first three areas (1-3) of Figure 5, which means that these tools had no influence on the identification of the platform configurations. All companies provide (1) PIM software in the form of E-Mail, network directories and shared notebooks. The software category of (2) Intranets (DXP/CSP) mostly contains complementary stand-alone tools. All companies have multiple (independent and redundant) (3) communication tools

such as chat and video conferencing. Looking back at Figure 2 (Areas of Collaborative Work), it is noticeable that the commonalities on the left side of the figure are for 1:n activities such as "top-down informing" and "content provision and preservation".

The differences in the configurations, however, can be identified mostly in the core areas (5-8) of multilateral joint work, the (5) *communication* among employees, the (6) *cooperative work* on documents, the (7) *coordination* of work and the joint (8) *combination and enriching of documents*.

Table 1 contains an analysis of the locations of the functional modules in the three case companies. Numbers larger than 1 indicate redundant functionality. Characteristically for the *Concentration Approach*, the case company has built its ECP almost entirely on the core ECS. They provide their employees with only a few additional software components specialised on communication and task management. In the example of the *Diversity Approach*, identical functionality is provided by all 3 systems/suites (files, tasks and wikis). The case company in the example for the *Dual Core Approach* provides two suites with some overlapping functionality and some functional areas that are not supported (surveys and forums).

Table 1. Location of functional modules in the ECPs of the three case companies

<i>Module</i>	<i>Concentration</i>	<i>Diversity</i>	<i>Dual Core</i>
Files	(2) ECS (files), network directories	(4) all 3 ECS/suites, network directories	(3) Suite 1 (OneDrive, SharePoint, Office 365)
Microblogs/Chat	(5) ECS (status update), Slack, Skype, Whatsapp, Sametime	(3) ECS (status update), Suite 1 (Teams, Skype)	(3) Suite 1 (Yammer, Skype, Teams)
Video Conf.	(2) Zoom, Skype	(2) Suite 1 (Teams, Skype)	(2) Suite 1 (Teams, Skype)
Surveys	(0) -	(1) ECS (Surveys)	(0) -
Tasks	(2) ECS (activities), Trello	(4) ECS (activities), Suite 1 (Planner, Todo) Suite 2 (Jira)	(1) Suite 2 (Jira)
Wiki	(1) ECS (wiki)	(3) all 3 collections	(2) Suite 1 (Teams wiki), Suite 2 (Confluence)
Forum	(2) ECS (forum), Notes Forums	(1) ECS (forum)	(0) -
ESN	(1) ECS	(1) ECS	(1) MS Suite (Yammer)

The findings reveal that there are multiple emerging platform strategies and designs. The three case examples presented above show distinct approaches, one of more strict control over the number of tools in use by focusing on the functionality of the core ECS (*Concentration*), a second more open, offering multiple tools to support the same functionality and thus giving the employees more flexibility but also the burden of choice (*Diversity*). The third approach (*Dual Core*) combines two collections for the necessary range of tools to create the Enterprise Collaboration platform which also creates some redundancy.

In summary, our research identified that Enterprise Collaboration Platforms are built around one or multiple *core* products that are complemented by multiple additional tools and applications to provide the desired comprehensive support for collaborative work.

Conclusions and Outlook

In this research note attention is focused on the large-scale provision of collaboration software in organisation to provide a stable technology environment to support a wide range of collaborative work situations. To date, limited work has examined the enterprise level provision of software tools. Existing studies are frequently focused at the micro-level largely limited to studies of a single type of social software (e.g. blogs, wikis) (Holtzblatt et al., 2010; Richter & Riemer, 2013) or to a specific type of collaborative activity (e.g. knowledge sharing, expert search) (Hacker et al., 2017; Mäntymäki & Riemer, 2016) and lack the scale and scope to address the wider, meso/organisational and platform level transformations to the organisation-wide support of collaborative work.

In this study we intentionally focus on the enterprise-wide provision of collaboration tools in order to examine the types of collaboration platform currently being assembled. Our findings confirm that currently, no single collaboration system provides all the functionality needed to support distributed collaborative work, thus requiring organisations to assemble their collaboration platform by selecting and combining additional tools and systems. The study examines these portfolios of tools and provides a classification of areas of collaboration support (ArCoW) and a catalogue of software with functionality in these areas. Using these insights, a survey to examine the portfolios of tools in 23 organisations was conducted. The data for each organisation was then analysed and visualised and used to identify the three dominant collaboration platform configurations described above. The visual representation of collaboration platform structures reveals the diversity and complexity of the collaboration technologies supported and will enable us to further examine enterprise-level collaboration support and to consider the design requirements for a more integrated platform for collaborative working.

The current study provides insights into which technologies are in use; the questions of how these platforms are planned and designed remain open. Based on this, the next phases of our work are focused on investigating the platform design processes to understand how these enterprise collaboration platforms are being designed (for example through intentional centralised planning, the situated requirements of individual work groups or both) and how they are evolving over time. For example, is functional redundancy removed (or does it increase), do the portfolios of tools continually change or do they stabilise?

References

- Bafoutsou, G. and Mentzas, G. (2002): 'Review and functional classification of collaborative systems', *International Journal of Information Management*, vol. 22, pp. 281–305.
- Creswell, J. W. and Plano Clark, V. L. (2018): *Designing and conducting mixed methods research*, (3rd ed.) SAGE Publications Inc, Thousand Oaks, CA.
- Ellis, C. A., Gibbs, S. J. and Rein, G. (1991): 'Groupware: some issues and experiences', *Communications of the ACM*, vol. 34, no. 1, pp. 39–58.
- Gartner (2021): 'Gartner Forecasts Worldwide Social Software and Collaboration Market to Grow 17% in 2021', retrieved from: <https://www.gartner.com/en/newsroom/press-releases/2021-03-23-gartner-forecasts-worldwide-social-software-and-collaboration-market-to-grow-17-percent-in-2021>.
- Gartner (2022a): 'What are Digital Experience Platforms (DXP) Reviews and Ratings', retrieved from: <https://www.gartner.com/reviews/market/digital-experience-platforms>.
- Gartner (2022b): 'Gartner Glossary: Identity and Access Management (IAM)', retrieved from: <https://www.gartner.com/en/information-technology/glossary/identity-and-access-management-iam>.
- Gartner (2022c): 'Content Services Platforms (CSP) Reviews and Ratings', retrieved from: <https://www.gartner.com/reviews/market/content-services-platforms>.
- Gratton, L. (2021): 'How to do hybrid right', *Harvard Business Review*, vol. 99, no. May-June.
- Hacker, J., Bodendorf, F. and Lorenz, P. (2017): 'Helper, Sharer or Seeker? – A Concept to Determine Knowledge Worker Roles in Enterprise Social Networks', *13th International Conference on Wirtschaftsinformatik*, pp. 668–682.
- Hausmann, V. and Williams, S. P. (2016): 'Issues for the long-term management of Social Business Documents', *International Journal of Information Systems and Project Management*, vol. 4, no. 3, pp. 45–61.
- Holtzblatt, L. J., Damianos, L. E. and Weiss, D. (2010): 'Factors impeding wiki use in the enterprise: A case study', *Proceedings of the 28th ACM Conference on Human Factors in Computing Systems*, pp. 4661–4676.
- Lee, C. P. and Paine, D. (2015): 'From The Matrix to a Model of Coordinated Action (MoCA)', *18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW 2015)*, pp. 179–194.
- Leonardi, P. M., Huysman, M. and Steinfield, C. (2013): 'Enterprise Social Media: Definition, History, and Prospects for the Study of Social Technologies in Organizations', *Journal of Computer-Mediated Communication*, vol. 19, no. 1, pp. 1–19.
- Mäntymäki, M. and Riemer, K. (2016): 'Enterprise social networking: A knowledge management perspective', *International Journal of Information Management*, vol. 36, no. 6, pp. 1042–1052.
- McKinsey Global Institute (2021): *The postpandemic economy: The future of work after COVID-19*, MGI Report February 2021.
- Richter, A. and Riemer, K. (2013): 'Malleable end-user software', *Business and Information Systems Engineering*, vol. 5, no. 3, pp. 195–197.
- Riemer, K. (2007): 'The Market for E-Collaboration Systems - Identification of System Classes Using Cluster Analysis', *European Conference on Information Systems 2007*, pp. 346–357.
- Riemer, K., Overfeld, P., Scifleet, P. and Richter, A. (2012): 'Eliciting the anatomy of technology appropriation processes: A case study in enterprise social media', *European Conference on Information Systems*.

- Schubert, P. (2018): 'Joint Work and Information Sharing in the Modern Digital Workplace: How the Introduction of "Social" Features Shaped Enterprise Collaboration Systems', in K. Riemer, S. Schellhammer and M. Meinert (eds.): *Collaboration in the Digital Age: How Technology Enables Individuals, Teams and Businesses*, pp. 45–59, Springer, Berlin, Heidelberg.
- Schubert, P. and Williams, S. P. (2022): 'Enterprise Collaboration Platforms: An Empirical Study of Technology Support for Collaborative Work', *Procedia Computer Science*, vol. 196, pp. 305–313.
- Wehner, B., Ritter, C. and Leist, S. (2016): 'Enterprise social networks: A literature review and research agenda', *Computer Networks*, vol. 114, pp. 125–142.
- Williams, S. P. (2010): *Enterprise 2.0 and Collaborative Technologies*, (Issue May 2010) Working Report of the Research Group Business Software, May 2010, University of Koblenz-Landau, Koblenz, Germany.
- Williams, S. P. and Schubert, P. (2011, June): 'An Empirical Study of Enterprise 2.0 in Context', *24th International Bled Conference*.
- Williams, S. P. and Schubert, P. (2017): 'Connecting Industry: Building and Sustaining a Practice-based Research Community', *50th Hawaii International Conference on System Sciences*, pp. 5400–5409.
- Williams, S. P. and Schubert, P. (2018): 'Designs for the Digital Workplace', *Procedia Computer Science*, vol. 138, pp. 478–485.
- World Economic Forum (2020): *Resetting the Future of Work Agenda: Disruption and Renewal in a Post-COVID World*, (Issue October) Cologne/Geneva, Switzerland.

Annex

Collaborative Technologies - Evaluation Tool based on the 8C Model for Enterprise Information Management

Scale:

0 - function not supported

1 - function supported

EVALUATION CRITERIA		TOOLS				Example functionality/typical functionality of:
		Req.	Tool1	Tool2	Tool...	
COMMUNICATION (TOTAL)		0	0	0	0	
1	Chat (text message)	0	0	0	0	e.g. SMS, WhatsApp
2	Microblogging	0	0	0	0	e.g. Twitter tweet, Yammer
3	Blogs	0	0	0	0	e.g. diaries, 1 to many
4	Voice message synchronous	0	0	0	0	e.g. Skype, VOIP
5	Voice message asynchronous	0	0	0	0	e.g. voice box ("answering machine")
6	Asynchronous sent (rich) text message	0	0	0	0	e.g. e-mail
7	Discussion forums	0	0	0	0	e.g. forum, newsgroups (hierarchical message structure: post and r
8	Notice boards	0	0	0	0	e.g. boards/pinboards (broadcast, posts, no responses)
9	Comments, annotations	0	0	0	0	e.g. comments on files
10	Video conferencing	0	0	0	0	e.g. Skype, Zoom
11	Unified Communication	0	0	0	0	e.g. information integration on different devices, e.g. same
12	Broadcast	0	0	0	0	e.g. webcast, podcast
13	### (other - please add any other criteria not already liste	0	0	0	0	
COOPERATION (TOTAL)		0	0	0	0	
1	Shared authoring	0	0	0	0	e.g. text editors, Wikis, whiteboard
2	Markup of changes (in a text)	0	0	0	0	e.g. revision marks
3	Screen sharing/shared desktop	0	0	0	0	e.g. with Teamviewer
4	Shared workspaces	0	0	0	0	e.g. a group space, team room, community
5	Workspace awareness	0	0	0	0	e.g. information on what is happening in the past, presence and
6	User profiles / social profiles	0	0	0	0	e.g. to show demographics, personal attributes, discovery of
7	Ratings, rankings	0	0	0	0	e.g. "useful", "not useful", "most frequently viewed" (shows first c
8	### (other - please add any other criteria not already liste	0	0	0	0	
CONTENT COMBINATION (TOTAL)		0	0	0	0	
1	Document management (document storage, archiving)	0	0	0	0	e.g. EDRMS type functionality (e.g. contracts, manuals)
2	Content management	0	0	0	0	e.g. CMS type functionality (e.g. web pages, images, descriptions)
3	Data aggregation (display what a user needs on start page	0	0	0	0	e.g. portals, corporate entry pages
4	Data integration	0	0	0	0	e.g. mashups, dashboards, widgets
5	Content collection	0	0	0	0	e.g. Wikis, glossary
6	Linking (e.g. hyperlinks)	0	0	0	0	e.g. hyperlink in posting
7	Pointers or references to content	0	0	0	0	e.g. bookmarks
8	Tagging, Folksonomies	0	0	0	0	e.g. classifying label (tag) on content items (e.g. posts)
9	Visualisation of tag usage	0	0	0	0	e.g. tag cloud or tag list
10	Collecting feedback	0	0	0	0	e.g. with surveys, ratings
11	Search	0	0	0	0	e.g. search in posts and documents
12	Content subscription	0	0	0	0	e.g. RSS feeds
13	### (other - please add any other criteria not already liste	0	0	0	0	
COORDINATION (TOTAL)		0	0	0	0	
1	User directories	0	0	0	0	e.g. LDAP, groups, distribution list, shared address book
2	Roles	0	0	0	0	e.g. organisational roles (e.g. manager, assistant) and access rights
3	Group calendar, deadline planning	0	0	0	0	e.g. joint view of appointments of multiple people
4	Resource planning	0	0	0	0	e.g. reservation of resources (rooms, projector, people)
5	Shared tasks	0	0	0	0	e.g. todo lists involving multiple people
6	Reminders, triggers, alerts	0	0	0	0	e.g. notification when response has been written or threshold rea
7	Workflow support	0	0	0	0	e.g. predefined flow of documents including authorisation (signed
8	Graphical flow	0	0	0	0	e.g. visualisation of a document flow and attached rules
9	Polls and voting	0	0	0	0	e.g. to find a decision in a team
10	Document and version control	0	0	0	0	e.g. automated versioning
11	Presence awareness	0	0	0	0	e.g. status icon in Skype
12	### (other - please add any other criteria not already liste	0	0	0	0	

Huh-Yoo, J., Montgomery, M., Matta, R., Kim, H., Daly, B. (2022): Toward Supporting the Mental Health of Underprivileged Youth Through Village-Driven Sociotechnical Systems. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Notes, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_n02

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_n02

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

Toward Supporting the Mental Health of Underprivileged Youth Through Village-Driven Sociotechnical Systems

Jina Huh-Yoo¹, Meredith Montgomery¹, Rishabh Matta¹, Heejun Kim³, Brian Daly²

¹Drexel University Department of Information Science, ²Drexel University Department of Psychological and Brain Sciences, ³University of North Texas Department of Information Science

jinahyoo@drexel.edu, mm5363@drexel.edu, rm3556@drexel.edu, heejun.kim@unt.edu, bpd36@drexel.edu

Abstract. Children and adolescents from socially and economically disadvantaged backgrounds are significantly more likely to suffer from mental health problems. Underprivileged youth experience various stressors, including racism and exposure to violence, leading to more violence. Despite the wide adoption of mobile health platforms among the target population, the literature lacks evidence on using mobile health to support underprivileged youth. The community leaders of urban neighborhoods came together to share thoughts on how mobile solutions can best support the mental health needs of underprivileged children and adolescents. Through interviews with pastors, principals, and non-profit organization leaders, we show what is needed is a socio-technical solution that engages the community to create an ecosystem of village-driven support for youth mental health. We found a disparity between the outcomes the community is interested in improving (e.g., resilience, personal agency, mentorship) versus the clinical outcomes usually measured in mental health mobile health applications. The community desire to work on “village-driven” and engagement-based approaches rather than the intervention of clinical/treatments more common in existing mobile health solutions. We discuss

implications for the requirements for village-driven socio-technical systems in supporting the mental health of underprivileged youth.

1. Introduction

Children and adolescents from socially and economically disadvantaged backgrounds are significantly more likely to suffer from mental health problems than their peers from affluent backgrounds (DeFosset et al. 2017). Despite these alarming mental health disparities, overwhelming evidence reveals that underprivileged youth and children of color in the U.S. are the least likely to receive quality mental health services (Ho et al. 2007; Planey et al. 2019). Known barriers include stigma and shame, lack of trust toward healthcare providers, cultural differences, religious preferences, poor access to care, and concerns about cost (Breland-Noble et al. 2011; Graves 2017; Lindsey et al. 2013; Weiler, Boat, and Haddock 2019). These patterns are evident in countries around the world, including the U.K. (Kirby, Wright, and Allgar 2020), Sweden (Kim and Hagquist 2018), the Netherlands (Weinberg et al. 2019), and Germany (Reiss 2013). Across the United States and Europe, the prevalence of mental health problems is the highest among minority youth living with economic instability (Alegria, Vallas, and Pumariega 2010; Reiss 2013). For example, in the United States, Black and African American people living below poverty are twice as likely to report severe psychological distress compared to those living over twice above the poverty level (National Center for Health Statistics 2019). In the U.K., Deighton and colleagues reported that the odds ratios for experiencing mental health problems among school-age children and adolescents were much higher based on ethnicity and socioeconomic status (Deighton et al. 2019).

More than a quarter of the Philadelphia population lives below the poverty line [11]. Philadelphia also suffers from a gun violence crisis, where there have been 3,884 shooting incidences and 2,222 shooting victims in 2020, a staggering 66.1% and 53.7% increase from 2019 (“Crime Mapper: Philly Police,” 2022). More than 100 victims were juveniles [12]. The city is the 4th most segregated city in the U.S. Philadelphia has over 100 separate neighborhoods with their own historical and residential characteristics. Those neighborhoods with a lack of community cohesion or experiencing transitions have consistently exhibited higher rates of gun violence than other neighborhoods [15–17]. Such disproportionate amounts of exposure to gun violence in specific neighborhoods adversely affect mental health (Smith et al. 2020).

Participating in community-building activities, mentorship, and community-led interventions can positively impact the mental health and emotional well-being of those who provide and receive the services (Castillo et al. 2019). Supportive social networks, existing relationships with therapists, and referrals and mandates by

parents and community-based gatekeepers facilitate seeking mental health care (Bannon Jr et al. 2008; Lindsey et al. 2006; Samuel 2015; Lindsey et al. 2013).

We conducted interviews with religious leaders, leaders of non-profit organizations for youth, teachers, and principals of K-12 schools to better understand how to best support the mental health needs of youth. The collaborative goal as a community was to plan, design, and develop a novel mobile solution that youth and people in the neighborhoods can use to improve the mental health of underprivileged youth. We present existing systemic challenges in supporting youth mental health and highlight the lack of research on how mobile solutions can effectively support youth mental health. We present “Village-driven support” as a viable solution to this challenging societal challenge and discuss the requirements for setting up a village-driven mobile technology to support the mental health of underprivileged youth.

2. Background

Promoting positive outcomes and support for healthy behaviors for youth requires access to a variety of assets and protective factors across the family, school, and community system (Henderson, DeCuir-Gunby, and Gill 2016; Reupert 2017). Neighborhood environments play a crucial role in mental health through experiences of violence, social connectedness, and socioeconomic conditions (Reupert 2017; Hill and Maimon 2013; Cox et al. 2017; Stockdale et al. 2007). While evidence has emerged around neighborhood factors impacting youth mental health, little work includes the dynamic nature of neighborhoods and community-building activities or investigates community-based social interactions on mobile applications.

A systematic review revealed several essential barriers to mental health help-seeking among Black youth and their families, including negative experiences with clinicians, provider mistrust, perceptions that treatments were ineffective, or that health professionals did not respond to their needs (Mukolo and Heflinger 2011; Murry et al. 2011; Samuel 2015; Graves 2017). For instance, 31% of black mothers endorsed the following statement: “White professionals could not understand the problems of African American families” (Murry et al. 2011). Seventeen percent endorsed the item “I am suspicious that White professionals would not treat my child as well as s/he would treat a White child” (Murry et al. 2011). A focus group study of Black children from 6th to 8th grade found “several adolescents raised the negative expectation or fear that their ‘parents might get taken away’ if they tell a therapist that they were hurt or abused” (Lindsey et al., 2013, p.113). While barriers exist in access to mental health services for Black youth, other approaches, such as community-based social support activities, have positively affected their mental health and emotional well-being. This finding is in line with how facilitators to accessing mental health services among Black youth include social support, such

as religious leaders or past mental health service providers sharing the importance of early clinical intervention (Lindsey et al. 2010; Cauce et al. 2002).

A review of mental health apps suggested promise for feasibility and acceptability for marginalized and underprivileged populations, although large-scale efficacy testing and scaling potential are still lacking (Schueller et al. 2019; Naslund et al. 2020). A systematic review of mHealth studies for minority and low-income populations experiencing health disparities revealed very few applications beyond SMS text messaging (Schueller et al. 2019). Most mHealth solutions for mental health focus on individuals. Community-based mobile solutions are under-investigated despite their potential. Given the sparse evidence and support for mHealth apps with underprivileged youth, engaging the users and relevant communities earlier in the development process is critical to designing an acceptable novel solution.

Veinot et al. conducted an interview study to develop a trust-centered design framework that guides designing technological solutions for Black youth (Veinot et al. 2013). Trust was cited as a critical concern for Black youth. The authors concluded that we must move beyond the currently accepted trust-promoting strategies to reach underprivileged populations adequately. And it was posited that we should investigate the social context and align it with cultural trust patterns in personal, group, institutional, and technological realms.

Harrington and Dillahunt used community-based participatory design methods to engage youth in marginalized communities in speculative design (Harrington and Dillahunt 2021; Bray and Harrington 2021). They partnered with an existing organization already situated and invested in community engagement. Such co-design workshops took a collectivist approach and centered on community experiences to address inherent potential power imbalances among designers, researchers, and community members (Bray and Harrington 2021). More of this community-engaging, participatory work will be critical to understanding and supporting the population and their needs and rethinking how existing mobile solutions can or cannot help with what the community is experiencing.

3. Methods

Our team consisted of researchers, students, and the community advisory board (CAB). The CAB reached out to the lead researcher of this team during the COVID Pandemic in the summer of 2020 to seek a mobile solution that would be amenable and acceptable to youth. CAB's goal would be to help prevent the rising violence and the disintegration of the city's neighborhoods. The CAB consists of pastors and reverends of several churches in the city. They also work as college counselors and voluntarily advocate for youth and older adults who need support. Notably, the project was initiated by the CAB. The CAB has continued to direct the research questions (below) and provide access to recruiting the community's leaders who advocate for underprivileged youth.

RQ1. What are the current challenges and solutions to supporting the mental health of underprivileged youth in Philadelphia?

RQ2. What are the requirements for building a mobile solution to support youth mental health?

We investigated RQ1 to understand the gap and availability of the current resources in the school and the community. With the preliminary perspectives gathered from CAB members and the literature, we sought to elicit opinions from the interviewees regarding what kind of solutions seem viable to support the mental health of youth. Hearing from youth about their direct experiences was the next phase in the research. Given that youth is a vulnerable human subject population, the CAB and the research team agreed it is safe first to interview the leaders to make sure the research team understands the population and its potential challenges and risks in conducting research with the youth.

We recruited eight leader interviewees with the following criteria: (1) aged 18 and older, (2) have an interest in supporting youth, (3) speak English, and (4) live in Philadelphia. The CAB provided a list of interviewees to contact, and we sent a questionnaire that consisted of the informed consent form and the screening questionnaire. Recruiting was highly challenging because of the COVID pandemic and the potential participants' work pressures and time constraints. We set up video-conference meetings with the recruited participants, where we recorded the interview for transcription purposes. The video-conference tool automatically transcribed the interviews. Students of the team went through the recording, fixed any errors, and deidentified any personal identifiers. We then imported the deidentified transcripts to NVivo (Richards 1999) for thematic analysis.

Using open coding analysis, three researchers went through one transcript and developed emerging codes grounded in the data (Strauss and Corbin 1994). Then, two researchers separately coded another transcript and shared the independently coded codes to understand the similarities and differences between the researchers. Three researchers then developed a codebook and conventions for coding, where we allowed new codes to be developed as the researchers encountered them in new transcripts. The two researchers then split the rest of the transcripts to conduct a thematic analysis based on the agreed codebook, which consisted of four classifications on systemic challenges to youth's mental health, existing community-based solutions, the gap in existing solutions, and the community's recommended solutions. The three researchers then came together to go through the codes together and conducted affinity diagramming to analyze common patterns and themes across the codes. These themes will be introduced in the findings. Table I indicates the participant demographics. The recruitment and the interviews were conducted after receiving approval from the university's IRB #2010108219A001.

Table I. Participant demographics

Participant#	Gender	Ethnicity	Age	Role in the community
P1	Female	African-American	74	A retired teacher, school director
P2	Female	African-American	39	Principal (grades 5-12)
P3	Female	White	65	Director of a health organization
P4	Female	African-American	51	Police staff inspector
P5	Male	African-American	N/A	Director of a non-profit organization
P6	Female	African-American	44	Principal (grades K-5)
P7	Female	African-American	N/A	Principal (grades K-8)
P8	Female	African-American	83	Reverend, community advocate

4. Findings

“Even if you do have mom and dad and grandmother and grandfather involved, there's so much to life. There're so many things you have to deal with in life that you really do need a village to survive.” – P4

The community leaders attributed violence and mental health issues among the underprivileged youth in their community to systemic challenges children and families were facing. Because of its systemic nature, particular to this target population, existing solutions developed for the general public would not be something that would work. Instead, the leaders saw the ‘village’ as a solution to bringing support at the individual, family, and community levels. However, these activities were often unscalable or met with challenges in engaging and sustaining the support of stakeholders such as families, children, and volunteers. Given the prevalent use of mobile phones among youth, the leaders saw mobile technology as a potential solution to help fill those gaps in connecting people, increasing engagement, and reaching out to people with resources. The resulting benefits they wanted to see were to increase resilience, personal agency, and competency in the youth, rather than clinically oriented mental health outcomes.

In this section, we describe the systemic challenges to the mental health of underprivileged youth and what solutions the leaders perceived as helpful.

Systemic challenges to the youth’s mental health

The systemic challenges the leaders saw that attributed to the mental health problems of youth included racism, trauma associated with violence, drugs, food insecurity, and poverty. These findings highlight the challenging context of the mental health of this population and suggest the need for different approaches than those used for other populations who do not experience these systemic challenges.

For instance, P3 explained how systemic racism and societal norms made “the inner city and youth of color” feel like “you are less than and feel like you don't

matter.” The pandemic and recent injustice toward Black men aggravated these experiences:

“Especially at the time we're in right now, children are angry. They're angry because they're stuck in their houses, they're angry because of social justice issues, they're angry just to be angry, they're angry just because. There's so much.” – P2

As a result, P2 stated that students without coping skills exhibit signs of self-harm and struggle with depression. Furthermore, due to poverty, P6 shared that youth have experienced trauma from the loss of income, having to move multiple times in a single school year (e.g., called transients), being abused by parents, and witnessing their parents' abuse and witnessing a homicide. P4 also shared drug addiction in the community:

“Individuals are suffering from meth addiction, so obviously their children will be impacted by, you know, these decisions that they've made.” – P4

The environment in which the youth grow to suffer from food insecurity: “In [anon]'s community there's no supermarket, there is no place where people can go and give fresh produce. There are several corner stores or fast-food places, but you know healthy eating is an important part of a child's overall health.” (P6) P2 shared how poverty leads to mental health problems exhibited in the classroom:

“then you've got poverty now trickling into mental health because now that student is suffering from anxiety, and then they're exhibiting that in their classrooms. You know, failing to pay attention, distractions in the classroom, anger. So I would definitely say poverty and mental health are the two largest struggles I'm currently facing. The students who I used to lead at a different school, I believe they're facing that, but they're also facing a lot of violence that is like in the city.” – P6

P8 shared how youth “truly think that they don't think they're gonna live long” and that this leads to the youth’s inability to plan their career long-term. P8 attributed this mindset to witnessing homicides and gun violence daily:

“I mean, if you're involved in a neighborhood where people getting killed five and six or day and people are shooting into neighborhoods, if you're having a 4th of July outing and somebody just ran, we come and shoot, that's kind of depressing and discouraging. They think that, oh, I could die any day. So I think that that deters their longevity of planning. In terms of a career.” – P8

Due to its systemic nature, entirely removing these societal challenges would be difficult. Instead, a solution to this challenge would be to find ways youth can build resilience, personal agency, and community support through education, activities, and mentorship.

Approaches to the systemic challenges

The leaders saw solutions to the systemic challenges as support driven by the village. For example, thinking of the community as whole, engaging family members, and working directly with youth personalized to their interests and contexts.

At the community level: Village-driven support

P4, who has a doctorate and is the city's Police Staff Inspector, attributed family and the community as what helped her to overcome her personal experiences of systemic racism, violence, and poverty:

"I witnessed violence. Individuals being shot. Individuals being stabbed. And the only thing that, although I felt traumatized in that instant, was my mother and my grandmother and my stepfather, we would go to church. The pastor would talk about those incidents. I know it sounds maybe a little strange, but I just never felt not protected. And those individuals who I was around, who, if they had a mother that was a substance abuser or father who wasn't around or a substance abuser or— there was always someone in that community in that village to look after them always." – P4

The leaders saw the community work as a 'village,' where the village becomes a resource for families. Being a resource required the community to be closely knit, where the members—adults and children—work together to support each other. Community members should act as support when a child's family cannot do so.

Being a resource included training people for peer mediation and conflict resolution (P1, P4), talking to youth about social issues that impact them (P4, P5), and being an outlet for aggression through arts and sports (P1, P3, P5). The key to these activities was that the community members were constantly engaged with the community. Children would gain the trusted resources they know others have tried (P2). When children feel equal to others, they believe it will help children become less violent (P5).

At the family level: Engaging parents

As the community builds cohesion, the community can target the families and individuals through various forms of mentorship activities between adults and children and among peers of children.

For a child to build good character, the leaders believed providing an environment with family values and high expectations and fostering positive engagement with family was a way to reduce violence, even within the family, such as domestic violence. The community's support towards families would create a suitable environment for children to feel protected and feel included. To foster this environment, the community should first help parents understand their child's situation if their children are at risk, provide opportunities to observe the child's behavior pattern, and assist the communication between parent and child. This activity should take a full-spectrum approach by coordinating teachers, counselors, and parents.

Furthermore, bringing human or technological resources to educate parents will be critical to help parents return to school or the workforce. It was essential to "get the conversation started" for parents, especially men, about their role in the family dynamic. The leaders shared using role-playing as a teaching method to engage parents further, talking about where they are, acknowledging their struggles, and doing things for the families. The leaders also believed pastors or community

ambassadors talking directly with the parents would help families feel comfortable and encouraged.

At the youth level: Finding interests and issues of scalability

Providing support at the level of families and directly to youth was critical. Examples included fostering talents and mentorship. However, these solutions are often met with a lack of resources and scalability challenges.

P5, who has run a non-profit organization to support underprivileged youth for 20 years, shared how community engagement reduced crime in the neighborhood they focused on:

“The portion of which where [organization name] started, within about of six-block radius there's not a whole lot of serious crime. About two years ago, the police district captain even drafted a letter that there's not a lot of violent crime in our area. And one of the reasons is because we're out there all the time. We don't really have a Community Center, but we have a schoolyard, we have a basketball program out there every summer. But even during the basketball program, it's the mentorship that kids are not allowed to curse. I don't care; you know how old you are. If you coach coaches, you're not allowed to curse and argue and stuff like that either; if not, our children can hold you accountable, we gotta do push-ups, for example.” - P5

P6 also believed in the importance of extracurricular activities that would attract youth to build on their talents to help them plan long-term career goals. But these resources are lacking in the community. P6 shared existing gaps in building the youth up based on their talents:

“So many black and brown students come to school with so many talents, but unfortunately, we have not tapped into those talents and use their talents as leverage [to help them stay in school and build career]. They're coming to school because they have basketball at the end of the day, or because they love playing flute, or because they have an art Program. You're going to send a child to then a middle school experience that might not have those same types of things or programs that you've had at the elementary school.” – P6

P3 similarly shared using sports to provide an outlet for talking with youth:

“For example, through a community center, build a new set of friends who are struggling with the same things instead of getting the aggression out by beating the crap out of somebody, you can get your aggression out by playing hoops and then have a conversation about how you feel, why you feel when you're not so angry.” - P3

However, finding volunteers who can steadily contribute to after school activities have been challenging for P7 because of the lack of resources:

“I steer away from volunteers unless they're truly able to commit to a certain time every day, and it's hard to. Everybody has a life to live, and it's tough trying to depend on that when you have children sitting there [waiting for the volunteer to show up].” – P7

The resources to help the youth in overcoming the obstacles in their schools and communities, such as counseling or extracurricular activities that would engage children to stay in school, are scarce and not scalable:

“It's mind-boggling that a seven-year-old child could have witnessed a homicide and has never had any type of counseling.” - P6

There is a lack of staff training in schools and community organizations, where the staff may “put their mental health in jeopardy trying to assist a student” (P2). Staff who assist students may not know where to find resources, and they end up all seeking help from the leaders:

“I’m their resource. And then me as a principal. It’s like this trickling effect; if I don’t have the resources to give them (40 staff members), they don’t have the resources to give the kids.” – P2

P2 continued to share the problem of lack of scalability:

“There are only one or two people to support a district of 200 and some other schools. Even if there’s five to support 200 schools, that’s not realistic, especially when teachers really need that training now and then they need a refresher at the beginning of the year.” – P2

Village-driven solutions to supporting youth’s mental health

The leaders commonly discussed mentorship and community-based activities as feasible and sustainable community-driven solutions to supporting the youth’s mental health. The leaders discussed finding the right mentor and what outcomes they would be looking for in supporting youth-centered, community-based activities.

Finding mentorship

The leaders considered mentor fit as a critical component in successful mentorship. The leaders discussed youth needing role models that might not be readily available for the child, such as a mother figure, a celebrity figure, or an older sibling (P2, P4, P5). Children should have common interests with the mentor (P2, P5) while having academic-focused mentorship with some discipline and more structure (P5). At the same time, peer mentors came up as potentially a better mentoring solution (P4, P5, P7) than older mentors because they respect the opinion of one another (P4). Mentors can have the skillsets and interests useful for youth for a career (P6) and those who excel and provide guidance (P2). P2 shared struggling youth should be paired with a mental-health competent mentor:

“There are some real self-harm issues. [It would be helpful to have] someone who may have been through depression and worked through, and someone who understands what isolation looks like but know how to come out of it. Someone who can also help them see the brighter side of things” – P2

The leaders also shared what not to do, which is to “come on strong right away to turn the child off or frighten them” (P1), not always pair like-minded individuals but add contrast (P4), and find mentors who are committed (P5). P1 shared that solid discipline is not good because it brings trauma (P1). Leaders mentioned the importance of asking the child who they are looking for and focusing on their interests, not what they’ve done (P1, P2, P3): “Ask them what they want. Don’t tell them what you think they want.”

Community-driven activities as part of the mentorship

The leaders shared various activities the community can do as part of the mentorship in supporting the youth's mental health. Leaders shared activities and clubs youth will desire, including basketball and football camps (P3, P5), video games (P1), workshops to discuss topics and issues or learn about interests like trades, business building, stage production, or fashion show (P1, P4, P5, P8), or to take field trips (P4). For instance, P3 shared that they are trying to use comics to teach about HIV and its danger to the youth. To enable self-driven and peer mentorship, P5 shared youth hosting a talk show as an idea he had already started implementing.

During these fun moments, academics and religion can be integrated, start conversations (P4), and bring opportunities to discuss incidents using the bible to help children cope (P4). These activities would give children an outlet for their struggles (P3): "get your aggression out by playing hoops." Youth can "build a new set of friends struggling with the same things." Bringing athletes, celebrities, city constituents, and experts to visit and talk to the youth would also help:

"We had a former [player from a famous basketball team]. They came they talked to the kids we had. [anonymous], who was the former coach of [a university], for you know 36 years, talk to the College men and women." -P3

"We have a lot of partners, so one of our partners because through [a local hospital's] healthcare network, so we had a guest come on and talk about the importance of physical activity" – P4

However, these resources are not always available to the youth, and appropriate training would be required to ensure the safety of everyone involved in establishing community-based mentorship relationships.

Community-based outcomes of interest

As a result of the community-based support, the leaders were not necessarily looking for clinically oriented mental health outcomes. Instead, they were interested in the outcomes integrated into the youth's career and life skills founded upon "resilience" and "personal agency."

The leaders emphasized "self-worth" as a critical element (P1, P2, P3, P4, P6). Academics are important, but P4 believed "adjustment" drives success in college enabled through resilience. The leaders shared how abused children could become resilient through encouragement, love, and praise by the community, for instance, despite their upbringing.

The personal agency would be built through uncovering gifts, talents, and abilities (P1). The goal is to stay in school, graduate from high school, pursue higher learning, excel in their careers, and become leaders (P1, P4, P5, P6, P8):

"[The goal would be] to have a game that reversed being a violent person to being an angelic person, so to speak, the person that goes around and spread great and goodness into the Community." – P1

The key was to foster children who would grow up to return what they received from the community and “teach and train other children” (P1, P7) once they grow and build their careers. P8 shares good examples of such community-based ecosystem that the community was able to build:

“[One struggling child] have gone on to college to be achievers. They’re in school now, and some of them have turned their lives around. They’ve given up being social marijuana smokers and now have active job careers, and so the work that we have done has proved fruitful.” -P8

The findings showed what the community needed to support the youth: community-driven mentorship. Building on the needs assessment, we discuss the implications of building socio-technical solutions to address these gaps.

5. Discussion

Systematic reviews of mobile health solutions for supporting mental health show very little evidence for those from underprivileged backgrounds who experience economic constraints, social stigma, and racism. Most mental health supporting systems have focused on clinically oriented domains specific to diagnoses, such as depression, anxiety, or substance abuse, and their direct outcomes, such as scales to measure reductions in symptoms of depression, physiological changes, or craving levels. However, the literature highlights that existing clinical approaches known to be effective for addressing mental health disorders lack evidence on whether the results generalize to minority populations. For instance, cognitive behavioral therapy is a well-known approach that has demonstrated effectiveness in many areas, such as post-traumatic stress disorder (Mendes et al. 2008), depression (Huguet et al. 2016), and insomnia (Trauer et al. 2015). At the same time, this approach has rarely been tested with low-literacy populations (Kuhada et al., 2011)

While these clinical outcomes may support measuring mental health status, the community instead focuses on the direct but long-term outcomes that they can observe as helpful for the individual and the community at the macro level. The community wants to create an ecosystem of community-driven, positive, and encouraging atmosphere for the children to grow in, regardless of their specific family or the societal and economic context that systemically hinders their positive growth.

Research shows the importance of being part of a community and participating in community-building activities. For example, community-led interventions can positively affect the mental health and emotional well-being of those who deliver and receive the services (Castillo et al. 2019). Such activities can increase a sense of belonging and social connectedness (Castillo et al. 2019). Furthermore, literature from the past 25 years indicates that appropriately matched mentors and youth can be effective in facilitating a range of positive developments and outcomes for many young people, including behavioral, social, health-related (e.g., emotional well-being), interpersonal, motivational, and academic outcomes (DuBois et al. 2011).

Our findings pointed to the community's need for a solid community-based mentorship program. Rhodes proposed that mentoring affects youth through three interrelated processes: (1) enhancing their social relationships and emotional well-being; (2) improving their cognitive skills through instruction and conversation; and (3) promoting positive identity development through meaningful connection to role models and advocates (Rhodes et al. 2006). In understanding when the mentor-mentee relationship is effective, little is known about the process by which matching characteristics influence the outcome of the match. Some data supports more enduring relationships and better outcomes when mentees and mentors have a voice and choice in selecting their partner (Rollin et al. 2003). The similarity of interests between mentor and mentee is another factor that has resulted in enhanced program outcomes as assessed by the length and quality of the relationship (Weiler, Boat, and Haddock 2019). Given the scattered evidence on how mentoring matches work, any sub-populations, especially with cultural sensitivity, will need additional research into what factors affect successful mentoring activities, especially for populations who lack evidence for important preferences.

Our findings uncovered what the community needs in selecting a mentor and what activities mentorship can support. The leaders have already been working hard to enable this ecosystem of village-driven support—trying to bring youth to after-school activities, connecting them with positive influences, and exposing them to educational environments. However, the leaders uniformly shared the lack of resources to scale these efforts that the community volunteers entirely drive.

The role of mobile technology for village-driven, the socio-ecological solution here, would be to connect people, expose them to positive environments, and engage stakeholders at all levels: youth, family (parents and guardians), staff of organizations involving youth, and community members and neighborhoods. Instead of approaching the gap with existing notions of mental health mobile health apps, the outcomes should focus on the factors that further sustain the community's ecosystem of social support and mentorship. The key theme of the support was connecting humans—connecting with people for positive influence. The technology solutions should not attempt to replace this social relationship but augment further connections that would otherwise have been difficult to achieve due to the limitations of physical space and the isolated nature of the pandemic. Youth will need engaging, fun, and creative ways to start conversations about their career, and interests, and share any challenges. Appropriate preventative help can be provided before the problem becomes more severe and difficult to remediate outcomes. Some concrete design requirements include, but are not limited to:

Recruit influential mentors with scalable and feasible solutions: The COVID-19 pandemic facilitated a wide adoption of mobile solutions to social interaction (e.g., mobile-based video conferencing tools, an app for voice-based communities), even among underprivileged communities (Al-Marroof et al.

2020; Clipper 2020; Finkelstein, Cha, and Wood 2012; Dey, Al-Karaghoul, and Muhammad 2020). Broad adoption of various asynchronous and synchronous social systems easily accessible through mobile phones provides opportunities for recruiting influential people to support the youth in creative ways. Examples of these high-impact people include celebrities, sports stars, peers, or former community members. These mentors can participate in the mentorship activities at various levels of engagement, from a recorded video to video-conferencing sessions or asynchronous chat sessions. For instance, celebrities can be available in an interactive system, such as chatbots, informing a celebrity's stories about overcoming adverse times. Youth can interact with this chatbot and ask questions, for instance.

Match mentors with youth based on user-informed algorithms: Technology can extract preferences from volunteers and youth in creative ways (e.g., have them write diaries, stories, or poems and use that to develop common interest points; or answer questions on interests and hobbies). Like dating apps, mentors and youth can state preferences for who they would like to be matched with, and the system will generate algorithms to find the best match. However, experts should closely monitor and moderate the process to ensure safety. It should also help facilitate communication between the mentor and youth and assist mentors with resources and training. The system should receive real-time user feedback, which can be immediately incorporated into the algorithms.

Foster location-based community-driven activities: Technology can facilitate online youth communities or a mix of community leaders and youth shaped by geographic neighborhoods. We should use theories on building online communities to build an interest-based community that would naturally evolve into neighborhood and inter-neighborhood activities that youth can enjoy. Effective promotion of accessible extracurricular activities based on location near the youth also will be helpful. Especially given the underprivileged population's lack of access to regular transportation, such location-based activities enabled by mobile solutions will be critical.

Promote immediate responsiveness: Technology should help youth immediately find solutions to any problems they feel stigmatized from sharing with others. If needed, the system could connect the youth with trained mentors or extracurricular activities that can address negative moments in real-time. For instance, youth can find peers to talk to who might have similar interests or mentors who built positive reputations by the peers for a talk to take aggression away in real-time. The technology should create immediate responsiveness to the requests, an essential part of the village-driven ecosystem supporting youth.

6. Conclusion

In this paper, we conducted interviews with community leaders advocating for underprivileged youth in a highly segregated city in the U.S. We identified village-

driven solutions as a sustainable, feasible solution to support youth at risk for mental health problems and exposure to community violence. We showed the unique outcomes the community was interested in, which vastly differ from existing mental health mobile health solutions. Furthermore, we highlighted the importance of using extracurricular activities to foster conversations among youth and mentors rather than framing the mentoring activities as mental health support. We also noted the importance of considering peers as mentors. Finally, we discussed implications for what socio-technical systems can do to fill the gap and what example systems might be helpful for youth in supporting and improving mental health.

7. Acknowledgments

We thank the CAB members, led by Rev. Mary Moore, Pastor Thomas, Rev. King and the iCAN Ministry. We also thank the community leaders who participated in the interviews.

8. References

- Alegria, Margarita, Melissa Vallas, and Andres J Pumariega. 2010. "Racial and Ethnic Disparities in Pediatric Mental Health." *Child and Adolescent Psychiatric Clinics* 19 (4): 759–74.
- Al-Marouf, Rana Saeed, Said A Salloum, Aboul Ella Hassanien, and Khaled Shaalan. 2020. "Fear from COVID-19 and Technology Adoption: The Impact of Google Meet during Coronavirus Pandemic." *Interactive Learning Environments*, 1–16.
- Bannon Jr, William M, Mary A Cavaleri, James Rodriguez, and Mary M McKay. 2008. "The Effect of Racial Socialization on Urban African American Use of Child Mental Health Services." *Social Work in Mental Health* 6 (4): 9–29.
- Bray, Kirsten, and Christina Harrington. 2021. "Speculative Blackness: Considering Afrofuturism in the Creation of Inclusive Speculative Design Probes." In *Designing Interactive Systems Conference 2021*, 1793–1806. New York, NY, USA: Association for Computing Machinery.
- Breland-Noble, Alfiere M, Carl C Bell, Antoinette Burriss, and AAKOMA Project Adult Advisory Board. 2011. "'Mama Just Won't Accept This': Adult Perspectives on Engaging Depressed African American Teens in Clinical Research and Treatment." *Journal of Clinical Psychology in Medical Settings* 18 (3): 225–34. <https://doi.org/10.1007/s10880-011-9235-6>.
- Castillo, Enrico G, Roya Ijadi-Maghsoodi, Sonya Shadravan, Elizabeth Moore, Michael O Mensah, Mary Docherty, Maria Gabriela Aguilera Nunez, et al. 2019. "Community Interventions to Promote Mental Health and Social Equity." *Current Psychiatry Reports* 21 (5): 1–14.
- Cauce, Ana Mari, Melanie Domenech-Rodríguez, Matthew Paradise, Bryan N Cochran, Jennifer Munyi Shea, Debra Srebnik, and Nazli Baydar. 2002. "Cultural and Contextual Influences in Mental Health Help Seeking: A

- Focus on Ethnic Minority Youth.” *Journal of Consulting and Clinical Psychology* 70 (1): 44.
- Clipper, Bonnie. 2020. “The Influence of the COVID-19 Pandemic on Technology: Adoption in Health Care.” *Nurse Leader* 18 (5): 500–503.
- Cox, Daniel T C, Danielle F Shanahan, Hannah L Hudson, Kate E Plummer, Gavin M Siriwardena, Richard A Fuller, Karen Anderson, Steven Hancock, and Kevin J Gaston. 2017. “Doses of Neighborhood Nature: The Benefits for Mental Health of Living with Nature.” *BioScience* 67 (2): 147–55.
- “Crime Mapper: Philly Police.” 2022. <https://www.phillypolice.com/crime-maps-stats/>. Accessed on 5/10/2022.
- DeFosset, Amelia R, Lauren N Gase, Roya Ijadi-Maghsoodi, and Tony Kuo. 2017. “Youth Descriptions of Mental Health Needs and Experiences with School-Based Services: Identifying Ways to Meet the Needs of Underserved Adolescents.” *Journal of Health Care for the Poor and Underserved* 28 (3): 1191.
- Deighton, Jessica, Suzet Tanya Lereya, Polly Casey, Praveetha Patalay, Neil Humphrey, and Miranda Wolpert. 2019. “Prevalence of Mental Health Problems in Schools: Poverty and Other Risk Factors among 28 000 Adolescents in England.” *The British Journal of Psychiatry* 215 (3): 565–67.
- Dey, Bidit L, Wafi Al-Karaghoul, and Syed Sardar Muhammad. 2020. “Adoption, Adaptation, Use and Impact of Information Systems during Pandemic Time and beyond: Research and Managerial Implications.” *Information Systems Management* 37 (4): 298–302.
- DuBois, David L, Nelson Portillo, Jean E Rhodes, Naida Silverthorn, and Jeffrey C Valentine. 2011. “How Effective Are Mentoring Programs for Youth? A Systematic Assessment of the Evidence.” *Psychological Science in the Public Interest* 12 (2): 57–91.
- Finkelstein, Joseph, Eunme Cha, and Jeffrey Wood. 2012. “Use of Social Media for Health Education: An Example of Online Support Group for Individuals with Atrial Fibrillation,” no. PG-31-34: 31–34.
- Graves, Lorraine. 2017. “Filters of Influence: The Help-Seeking Process of African American Single Mothers Living in Poverty Seeking Mental Health Services for Their Children.” *Child & Youth Services* 38 (1): 69–90.
- Harrington, Christina, and Tawanna R Dillahunt. 2021. “Eliciting Tech Futures Among Black Young Adults: A Case Study of Remote Speculative Co-Design.” In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–15.
- Henderson, Dawn X, Jessica DeCuir-Gunby, and Vandna Gill. 2016. “‘It Really Takes a Village’: A Socio-Ecological Model of Resilience for Prevention among Economically Disadvantaged Ethnic Minority Youth.” *The Journal of Primary Prevention* 37 (5): 469–85.
- Hill, Terrence D, and David Maimon. 2013. “Neighborhood Context and Mental Health.” In *Handbook of the Sociology of Mental Health*, 479–501. Springer.
- Ho, Judy, May Yeh, Kristen McCabe, and Richard L Hough. 2007. “Parental Cultural Affiliation and Youth Mental Health Service Use.” *Journal of Youth*

- and Adolescence* 36 (4): 529–42. <https://doi.org/10.1007/s10964-006-9114-x>.
- Huguet, Anna, Sanjay Rao, Patrick J McGrath, Lori Wozney, Mike Wheaton, Jill Conrod, and Sharlene Rozario. 2016. “A Systematic Review of Cognitive Behavioral Therapy and Behavioral Activation Apps for Depression.” *PloS One* 11 (5): e0154248.
- Kim, Yunhwan, and Curt Hagquist. 2018. “Mental Health Problems among Economically Disadvantaged Adolescents in an Increasingly Unequal Society: A Swedish Study Using Repeated Cross-Sectional Data from 1995 to 2011.” *SSM-Population Health* 6: 44–53.
- Kirby, Natalie, Barry Wright, and Victoria Allgar. 2020. “Child Mental Health and Resilience in the Context of Socioeconomic Disadvantage: Results from the Born in Bradford Cohort Study.” *European Child & Adolescent Psychiatry* 29 (4): 467–77.
- Lindsey, Michael A, Crystal L Barksdale, Sharon F Lambert, and Nicholas S Ialongo. 2010. “Social Network Influences on Service Use among Urban, African American Youth with Mental Health Problems.” *Journal of Adolescent Health* 47 (4): 367–73.
- Lindsey, Michael A, Kerri Chambers, Cara Pohle, Peggy Beall, and Alicia Lucksted. 2013. “Understanding the Behavioral Determinants of Mental Health Service Use by Urban, under-Resourced Black Youth: Adolescent and Caregiver Perspectives.” *Journal of Child and Family Studies* 22 (1): 107–21.
- Lindsey, Michael A, Wynne S Korr, Marina Broitman, Lee Bone, Alan Green, and Philip J Leaf. 2006. “Help-Seeking Behaviors and Depression among African American Adolescent Boys.” *Social Work* 51 (1): 49–58.
- Mendes, Deise D, Marcelo Feijó Mello, Paula Ventura, Cristiane de Medeiros Passarela, and Jair de Jesus Mari. 2008. “A Systematic Review on the Effectiveness of Cognitive Behavioral Therapy for Posttraumatic Stress Disorder.” *The International Journal of Psychiatry in Medicine* 38 (3): 241–59.
- Mukolo, Abraham, and Craig Anne Heflinger. 2011. “Rurality and African American Perspectives on Children’s Mental Health Services.” *Journal of Emotional and Behavioral Disorders* 19 (2): 83–97.
- Murry, Velma McBride, Craig Anne Heflinger, Sarah V Suiter, and Gene H Brody. 2011. “Examining Perceptions About Mental Health Care and Help-Seeking Among Rural African American Families of Adolescents.” *Journal of Youth and Adolescence* 40 (9): 1118–31. <https://doi.org/10.1007/s10964-010-9627-1>.
- Naslund, John A, Ameya Bondre, John Torous, and Kelly A Aschbrenner. 2020. “Social Media and Mental Health: Benefits, Risks, and Opportunities for Research and Practice.” *Journal of Technology in Behavioral Science* 5 (3): 245–57.
- National Center for Health Statistics. 2019. “Health.” Hyattsville, MD.
- Planey, Arrianna M., Shardé Mc Neil Smith, Stephanie Moore, and Taylor D. Walker. 2019. “Barriers and Facilitators to Mental Health Help-Seeking

- among African American Youth and Their Families: A Systematic Review Study.” *Children and Youth Services Review* 1 (101): 190–200.
<https://doi.org/10.1016/j.chidyouth.2019.04.001>.
- Reiss, Franziska. 2013. “Socioeconomic Inequalities and Mental Health Problems in Children and Adolescents: A Systematic Review.” *Social Science & Medicine* 90: 24–31.
- Reupert, Andrea. 2017. “A Socio-Ecological Framework for Mental Health and Well-Being.” Taylor & Francis.
- Rhodes, Jean E, Renée Spencer, Thomas E Keller, Belle Liang, and Gil Noam. 2006. “A Model for the Influence of Mentoring Relationships on Youth Development.” *Journal of Community Psychology* 34 (6): 691–707.
- Richards, Lyn. 1999. *Using NVivo in Qualitative Research*. Sage.
- Rollin, Stephen A, Cheryl Kaiser-Ulrey, Isabelle Potts, and Alia Haque Creason. 2003. “A School-Based Violence Prevention Model for at-Risk Eighth Grade Youth.” *Psychology in the Schools* 40 (4): 403–16.
- Samuel, Ignatius A. 2015. “Utilization of Mental Health Services among African-American Male Adolescents Released from Juvenile Detention: Examining Reasons for within-Group Disparities in Help-Seeking Behaviors.” *Child and Adolescent Social Work Journal* 32 (1): 33–43.
- Schueller, Stephen M, John F Hunter, Caroline Figueroa, and Adrian Aguilera. 2019. “Use of Digital Mental Health for Marginalized and Underserved Populations.” *Current Treatment Options in Psychiatry* 6 (3): 243–55.
- Smith, Melissa E, Tanya L Sharpe, Joseph Richardson, Rohini Pahwa, Dominique Smith, and Jordan DeVlyder. 2020. “The Impact of Exposure to Gun Violence Fatality on Mental Health Outcomes in Four Urban US Settings.” *Social Science & Medicine* 246: 112587.
- Stockdale, Susan E, Kenneth B Wells, Lingqi Tang, Thomas R Belin, Lily Zhang, and Cathy D Sherbourne. 2007. “The Importance of Social Context: Neighborhood Stressors, Stress-Buffering Mechanisms, and Alcohol, Drug, and Mental Health Disorders.” *Social Science & Medicine* 65 (9): 1867–81.
- Strauss, Anselm, and Juliet Corbin. 1994. “Grounded Theory Methodology: An Overview.” *Handbook of Qualitative Research* 273: 285.
- Trauer, James M, Mary Y Qian, Joseph S Doyle, Shantha M W Rajaratnam, and David Cunningham. 2015. “Cognitive Behavioral Therapy for Chronic Insomnia: A Systematic Review and Meta-Analysis.” *Annals of Internal Medicine* 163 (3): 191–204.
- Veinot, Tiffany C, Terrance R Campbell, Daniel J Kruger, and Alison Grodzinski. 2013. “A Question of Trust: User-Centered Design Requirements for an Informatics Intervention to Promote the Sexual Health of African-American Youth.” *Journal of the American Medical Informatics Association* 20 (4): 758–65. <https://doi.org/10.1136/amiajnl-2012-001361>.
- Weiler, Lindsey M, Ashley A Boat, and Shelley A Haddock. 2019. “Youth Risk and Mentoring Relationship Quality: The Moderating Effect of Program Experiences.” *American Journal of Community Psychology* 63 (1–2): 73–87.
- Weinberg, Dominic, Gonneke W J M Stevens, Elisa L Duinhof, and Catrin Finkenauer. 2019. “Adolescent Socioeconomic Status and Mental Health

Inequalities in the Netherlands, 2001–2017.” *International Journal of Environmental Research and Public Health* 16 (19): 3605.

Kristin Kaltenhäuser, Tijs Slaats, Thomas Gammeltoft-Hansen, Naja Holten Møller (2022): Deconstructing Gender in Asylum Categories: An Archival Perspective on a Practice with Limited Access. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Notes, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_n03

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_n03

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

Deconstructing Gender in Asylum Categories: An Archival Perspective on a Practice with Limited Access

Kristin Kaltenhäuser, Tijs Slaats, Thomas Gammeltoft-Hansen, Naja Holten Møller

University of Copenhagen

Contact Authors: krka@di.ku.dk, slaats@di.ku.dk, tgh@jur.ku.dk, naja@di.ku.dk

Abstract. Public authorities make decisions that greatly impact both citizens and non-citizens. Decision-making on asylum, which is regulated by international law but administered by states, in particular is characterised by a higher level of secrecy than other public services. The 1951 Refugee Convention defines refugeehood as the fear of being persecuted for reasons of race, religion, nationality, social group, or political opinion. Although fear of gender-related persecution was not included as one of the grounds meriting asylum, state practice means that it is today generally recognised as such. The United Nations Refugee Agency (UNHCR) recommends that states "ensure a gender-sensitive interpretation of the 1951 Refugee Convention." Using natural language processing (NLP) to analyse an open dataset of Danish asylum case summaries, we first identify five empirical categories connected to gender in the case summaries: 1) gender-related persecution, 2) LGBT 3) sexual conditions, 4) marital conditions and 5) other gender-related forms of persecution. Secondly, we illustrate the relationship between these gender-related categories and other categories/topics in asylum motives. Finally, we discuss how data science techniques can be applied to better understand complex, cooperative work practices in an area where access for researchers is limited, but archival data is available.

1 Introduction

Asylum decision-making generates unique data about casework, because it reflects how national and international law is put into practice. In this domain, casework is centered around an interview process aimed at assembling the required documentation to determine refugee status (Nielsen and Møller (2022)). In this paper we present findings from a study of gender-related categories in asylum case data. Gender-related persecution was addressed relatively late by international law compared to other categories of asylum motives (e.g. persecution for race, religion, nationality, social group, or political opinion). Even when it was formally introduced in 2002 by a UN guideline, gender continued to have a somewhat implicit status. It remains up to states to interpret and apply it (Byrne and Gammeltoft-Hansen (2020)). Our motivation for this study is to deconstruct the formal UN category of gender-related persecution and study its empirical categories. Using data science methods and applying an archival perspective (Thylstrup et al. (2021)), we investigate gender-related categories in an open dataset of 9,075 asylum case summaries handled by the Danish Refugee Appeals Board. We take an archival perspective in this study by tracing the origin of data structures and categories and raising questions about the power structures that shape them, in our case International law, national decision-making and archiving practice.

Asylum decision-making is a highly politicised and securitised area (Bigo (2014)) for which it is hard to negotiate access for observational studies. National authorities are often concerned for the applicants' safety and at the same time seek to avoid unwanted critique. That means that decision-making in this area is, from an access standpoint, much less public as compared to other types of administrative decision-making. Seaver studies a domain with similarly limited access, algorithms in an environment of corporate secrecy (Seaver (2017)), and introduces the term *scavenge*, to describe their research tactic. Scavenging denotes tracing cultural practices across heterogeneous locations to empirically characterise an inaccessible object of work (algorithms) without directly addressing it. Seaver emphasizes how an exploration across multiple locations can provide a better understanding of the persistent context of a practice and its work objects. The entry point for our scavenging is the Danish archive (database) that contains information about asylum-decision making (see Figure 1).

Following this line of argumentation, we ask: How are gender-related categories presented in the asylum case summaries and how can they be probed to understand the practice of asylum decision-making?

The study applies data science and natural language processing (NLP) techniques to conduct a category and topic analysis of asylum case summaries. Applying data science methods with an archival perspective is performative in the sense that it can bring about a new understanding of, in our case, gender-related categorisation practices in asylum. Our first contribution is identifying five empirical categories connected with gender in the asylum motives of the case

summaries: 1) gender-related persecution, 2) LGBT 3) sexual conditions, 4) marital conditions and 5) other gender-related persecution¹. Second, we illustrate the relationship between these gender-related categories and other asylum categories/topics, while showing how data science techniques used on archival datasets can serve as an entry point for studying practices in a context where access to the work domain is limited. We examine the empirical categories archived as part of a collaboration between the Danish national authorities and the Danish Refugee Appeals Board. When data is archived, power differences are present between the archiver who decides what is to be remembered and the individuals the records in the archive are about. While practices of categorisation and classification are a long-term interest of Computer-Supported Cooperative Work (CSCW) (Suchman (1993), Bowker and Star (1999), Møller and Bjørn (2011), Boyd and Crawford (2012), Pine and Liboiron (2015)), an archival perspective increases understanding of the contexts and power relationships that structure datasets. This can reveal something about underlying practices, in addition to observational studies which is a core strategy for understanding practices in CSCW (Randall et al. (2007)).

Engaging with a complex, collaborative work domain such as asylum, we argue, opens CSCW's long-standing interest in public sector decision-making for renewed considerations of how we as a research community can ensure that highly securitised areas of work are included in this strand of research. We as researchers can approach these sensitive areas with care for all stakeholders and act as intermediaries between interests of stakeholders with differential power, by increasing communication between the parties and making the power landscape visible.

2 Gender-related Persecution in International Asylum Law

While international law - notably the 1951 Refugee Convention - establishes a formal definition of what it means to be a refugee, states adapt and transform this category through both their national law and decision-making. The process of deciding asylum claims on a national level plays a major role in developing international refugee law. The lack of any dedicated court for refugees at the international level implies that states are the principal interpreters of international law in this area (Byrne and Gammeltoft-Hansen (2020)). The 1951 Refugee Convention is silent in terms of how states should design their asylum procedures. States implement diverse institutional and procedural frameworks for asylum decision-making, resulting in loosely coordinated, divergent practices across

¹ We note that the naming of the category *LGBT* in the dataset as a solitary initialism, lacking further description, is in itself questionable, since the asylum motive could in this case be read as *LGBT*, instead of *LGBT persecution*. Since part of the contribution of this study is the identification of empirical categories, we use the same naming, but raise this concern.

states. States maintain national databases of asylum decisions, establishing their own legal practices and categories on the basis of already decided asylum cases. In the case of Denmark, part of this database has been made publicly available. We analyse gender-related empirical categories in the public case data as a starting point for understanding the decision-making practices in this area.

The 1951 UN Refugee Convention describes a refugee as any person who:

owing to a well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion, is outside the country of his nationality and is unable or, owing to such fear, is unwilling to avail himself of the protection of that country (Article 1 A (2) Refugee Convention, UNHCR (1951))

Unlike some later human rights conventions, gender or sexual orientation is not a category explicitly listed in this formal definition. Since the 1980s, national asylum authorities have gradually come to recognise that gender may form the basis of an asylum claim; for instance, in some societies women can face repercussions for transgressing social mores, and sexual minorities can face widespread persecution. This position is backed by UNHCR guidelines, for example the 2002 guideline on gender-related persecution, which while not legally binding recommends states to "ensure a gender-sensitive interpretation of the 1951 Refugee Convention" as part of their national asylum procedures (UNHCR (2002)).

The UNHCR guidelines elaborate upon and specify the types of cases which belong within the category of gender-related persecution. As such, the UNHCR guidelines may be read as themselves constructing a category for seeking asylum, covering topics such as homosexuality, female genital mutilation, and so on. In the ensuing analysis, we systematically extract these topics, query the case summaries on this basis, and analyse their occurrence in the empirical categories applied by the Danish Refugee Appeals Board. The goal is to investigate the transformation of the formal category of gender-related persecution and its constituting topics when adopted in the national practice of asylum decision-making and reveal the topics that are omitted or articulated in the empirical categories that emerge in the Danish archive of asylum case summaries.

Gender has an implicit status in national asylum procedures, as it is not explicitly listed in the definition of refugeehood. According to the Convention, having fled one's country due to fear of persecution related to gender is not enough; one must have fled for one of five listed categories:

- race
- religion
- nationality
- membership of a particular social group
- political opinion

These categories are commonly referred to as *Convention grounds* in asylum law (see e.g. UNHCR (2002)) and provide an exhaustive list when deciding whether to grant a person refugee status based on the 1951 Refugee Convention.

For policy-makers to fit gender-related persecution into this framework, the introduction of a category hierarchy is required. Most commonly, gender is subsumed as a subcategory of *particular social group*, which has proven to be the most flexible category for types of asylum claims not originally foreseen when the Refugee Convention was drafted in 1951 (Hathaway and Foster (2003)). Legally, this subsumption is based on the argument that women and sexual minorities share the same type of core characteristics as regularly used to define other social groups. According to UNHCR:

It follows that sex can properly be within the ambit of the social group category, with women being a clear example of a social subset defined by innate and immutable characteristics, and who are frequently treated differently than men. Their characteristics also identify them as a group in society, subjecting them to different treatment and standards in some countries. Equally, this definition would encompass homosexuals, transsexuals, or transvestites. (UNHCR (2002))

Historically, this type of subsumption has equally led to categorisation struggles, with some asylum lawyers questioning whether *a particular social group* can meaningfully be broadened to include half the world's population. Even today, categorisation struggles continue within this area as states apply widely different approaches to defining the contours of who is at risk (Hathaway and Foster (2014)). As Spijkerboer notes, female asylum applicants thus tend to be constructed as a “double other” by Western countries, simultaneously singled out based on their gender and non-Western background (Spijkerboer (2000)). More fundamentally, the recognition of gender-related claims in this manner reproduces a historical category hierarchy, in which gender has to be constructed and defined by reference to the broader category of *membership of a particular social group* or other Convention grounds. As we will show, this has implications for how caseworkers subsequently approach and label asylum applicants fleeing gender-based persecution. In the ensuing analysis, we investigate the connection between gender and the formal list of Convention grounds, and how this category hierarchy is applied in Denmark.

3 Related Work

An archival perspective can support the rebalancing of datasets in favour of people that are displaced, prior research in archival studies suggest. For example Gilliland (2017) finds that official records and archives of states do not address the existential needs for documentation of non-citizens, such as asylum seekers, who

seek protection in a country, but do not enjoy the same rights as citizens. Cakici et al. (2020) investigate the relationship between citizenship and the data practices that shape how people can make claims to protection. Claiming data that is considered credible by decision-makers is not a straightforward process (Nielsen and Møller (2022)). We thus turn our attention towards related work on archival data and the associated knowledge work as arenas of power (Star and Strauss (2004)).

3.1 Archived Data as an Arena of Power

Increasingly, data in asylum decision-making is saved to databases (both national and international) that come to form an archive, a long-standing interest of CSCW (Ackerman and Malone (1990), Bannon and Kuutti (1996), Ackerman et al. (2013)). Recent research co-created, together with volunteers of social clinics, a digital archive of oral histories to generate a counter narrative to the official records of the healthcare system in Greece, serving both asylum seekers and other vulnerable groups (Vlachokyriakos et al. (2021)). The archive, as we learn from this research, may not simply be left alone for us to engage with it years later, but can play an active part in shaping practices. This counters a dominant understanding of the archive as a mere historical database and window into the past (Thylstrup (2022)).

Following Foucault, an archive can be an empirical site functioning as an administrative tool for data production about e.g. populations and nations, but also an analytical concept or a lens (Foucault (1972)). Power from this perspective is *generative* and understands communities as subjects through data, not simply as subjected to a power-over reality imposed onto them. Taking an archival perspective means paying close attention to various forms of power structures that shape archives (D'Ignazio and Klein (2020)) and thus become scrutable (Thylstrup (2022)). Amelia Acker in Thylstrup et al. (2021) suggests that investigating hidden origins of database structures (archives) and the motivation behind their implementation can provide insights into the power of cultural practices. She contends that categories of archives are continuously changing and subject to political and social change.

The archivists of the asylum decision-making process are the caseworkers (knowledge workers) who establish credibility, negotiate a status with a displaced person and construct them as an asylum seeker. Recognizing that categories have politics, as they are a fundamental device by which societies constitute their social order (Suchman (1993)), this form of knowledge production is a question of power, as specific categories are assigned to a person applying for asylum.

In a CSCW-context, we study practices of categorisation that come to form the origins of database structures (Møller and Bjørn (2011)). However, not all work practices are readily available for us to study in CSCW. The sensitive and highly politicised area of work in asylum decision-making is an example of a domain with limited access, where archived data can serve as an entry point for further study.

3.2 Data Shadows in Archives

Archived data casts *shadows* on topics that are omitted in the categories that make up the database structure and function as a gatekeeper for later access. People that are displaced rely on documentation and public records to establish credibility in order to be categorised as *persecuted* (Nielsen and Møller (2022)). However, these records are dispersed over several national and international archives, and people that are displaced often resort to "irregular forms and uses of records" to claim refugee status, e.g. washing out stamps in their passports that would make it impossible to enter certain countries or transmitting photographs of personal documents via mobile phones (Gilliland (2017)). Following Gilliland, publicly available and internationally accessible archives can play a critical role to address the needs of the displaced. In this sense, the public archive of asylum case summaries, we are analysing, may serve as a resource for asylum seekers and their advocates to navigate the Danish asylum decision-making process.

Data shadows are one example of an analytical concept that can make power differentials in datasets visible (Møller et al. (2021)). They appear when the realities of an actor with less power are not captured in a dataset. By inspecting data produced by asylum casework for data shadows instead of uncritically using it as ground truth for further processing (Muller et al. (2021), Aragon et al. (2022)), we make omitted and articulated topics visible and contestable.

Little research in CSCW has investigated how scholars can consider the relation between communities and the data is produced by and about them, to enable the construction of their own interpretations, as opposed to data technologies imposing a reality onto them (Irani et al. (2010), Taylor et al. (2015)). Caselli et al. (2021) argue for a participatory approach to NLP in which *language* is not considered as data, but as the product of people and advocate for involving producers of text in its sensemaking process (Ibid). We position our research in relation to debates in CSCW on changing the discourse around databased technologies that address how we respond to global connectivity and mobility (Irani et al. (2010)). Combining data science methods with an archival perspective allows us to scrutinize both the omitted and articulated gender-related categories in an archive of asylum cases, to study power structures that shape the national asylum decision-making process in Denmark.

4 Topic Analysis of Asylum Case Summaries

The dataset and archive we investigate makes summaries of asylum cases available for query on the website of the Danish Refugee Appeals Board. Figure 1 shows how categories are listed with check boxes so the user can search for entries associated with that category. In this case the categories take on the role of a gatekeeper to the archive. Topics that are articulated in the categories are easily accessible, while others that are not, require explicit searches or are otherwise forgotten.

The screenshot shows the website interface for the Danish Refugee Appeals Board. At the top, there is a navigation menu with options like 'Statistik og måltal', 'Nyheder', 'Kontakt', 'Sitemap', 'Links', and 'English'. Below this is a search bar with a 'Søg' button. The main content area is titled 'Praksis' and includes a sub-section 'Find praksismateriale'. This section has two dropdown menus: 'Vælg land fra liste:' (set to 'Alle lande') and 'Vælg afgørelsesår:' (set to 'Alle år'). Below these are two columns of checkboxes representing various asylum motives. The first column includes categories like 'Alfhængighedsforhold', 'Anden karsrelateret forfølgelse', 'Chikane', 'Familer med barn', 'Generelle forhold', 'Kriminelle forhold', 'LGBT', 'Modtageforhold', 'Overgreb', 'Privatretlig forhold', 'Religiøse forhold', 'Tilbageholdelse', 'Tortur', 'Uforholdsmæssig straf', and 'Ændrede forhold'. The second column includes categories like 'Agents of Persecution', 'Asylsagsproceduren', 'Etniske forhold', 'Familerelationer og ægteskabsliggende forhold', 'Helbredsmaessige forhold', 'Karsrelateret forfølgelse', 'Militære forhold', 'Nationalitet', 'Politiske forhold', 'Privatretlig konflikt', 'Seksuelle forhold', 'Tilknytt bistand oprørsgruppe', 'Udrejseforhold', 'Ægteskabelige forhold', and 'Øvrige modsætningsforhold til myndighederne'. A 'Søg' button is located at the bottom right of the category list. On the right side of the page, there is a 'Tips til søgning' section with instructions on how to use the search filters.

Figure 1. This Figure shows the interface to access the public dataset on the website of the Danish Refugee Appeals Board - <https://fln.dk/praksis> and the predefined categories that can be chosen to query the data.

4.1 Data Extraction and Analysis

Using techniques from NLP and a qualitative analysis of judicial text, we combine several analysis methods in this study. The data was scraped using the Python libraries *beautifulsoup4* and *Selenium* on the 19. October 2021 and yielded 9,075 case summaries. We performed the analysis of the category of gender-related persecution in the following steps:

1. After an initial exploratory analysis of asylum motives in the data and an investigation of the UNHCR definition of gender-related persecution, we matched and selected the categories in the Danish dataset with that definition for further analysis. The following are the five empirical categories:
 - gender-related persecution
 - LGBT
 - sexual conditions
 - marital conditions
 - other gender-related persecution

Following the UNHCR guidelines, claims based on sexual orientation contain a gender element. We therefore add the categories *LGBT* and *sexual*

conditions to our analysis. In an initial investigation, we also included the category of *abuse* as gender-related. After a qualitative study of a sample of the related case summaries, however, we discarded it. In this context, abuse mostly pertained to political and military contexts. For this study, we translated the categories into English. It is also important to note that in the Danish language, the same word is used for *sex* and *gender*.

2. We then analysed the other asylum motives with which gender-related categories are tagged. We first assessed if there are cases that only received a gender-related category and then calculated a co-occurrence matrix, as displayed in Table III. A co-occurrence matrix is a tool often used in NLP and image analysis. It represents the number of times terms appear in the same context with other terms (or pixels in the case of image analysis). In our case, we created a table in which each column is assigned to an empirical category pertaining to gender-related persecution in the Danish dataset. Each row represents one of the 29 asylum motive categories used by the Refugee Appeals Board.

The number in each cell of the table represents the number of cases to which a gender-related category is applied, together with another asylum motive category. The shading of the cell indicates the magnitude of the proportion of the number compared to the other categories that were tagged with it. We applied the shading for clarity to illustrate the composition of the co-tagged categories in the table.

3. We then systematically extracted topics from the UNHCR guidelines on gender-related persecution that are explicitly stated as claims related to gender. Table I provides a list of the extracted topics in the first column. We arrived at the list topics by identifying themes stated as examples relevant to gender-related persecution in the UN guideline. We decided not to include more general themes such as "transgression of social mores" (UNHCR (2002)), since it can include several topics, such as abortion, extramarital affairs, which sometimes overlap with other topics, which we included.
4. Finally, we conducted a topic analysis by querying the case summaries for words related to the extracted topics, see Table IV. For example, to detect the topic of *homosexuality*, we query the cases for the terms *lesbisk* and *homoseks*. To get from the UN topics to the Danish search terms that are applied in the case summaries, we performed a qualitative analysis on a sample of the cases (n=30) to select the keywords corresponding to the UN topics. This analysis showed a consistent use of the Danish terms as shown in Table I, except for transvestism, for which no cases were found. We employed manual stemming of the query words to reduce words to their basic form or stem. This technique is often applied in NLP to account for different forms of words in a text. Each case that contains any of these terms is counted as one occurrence. That means a case can contain several topics.

We then calculated the percentage composition of each category to illustrate which topics constitute a category.

Topic	Queried terms
Homosexuality	homoseks, lesbisk
Human trafficking	menneskehandel, traffick
Female genital mutilation	omskåret, omskæring
Forced marriage	tvangsgift
Forced prostitution	prostitution
Rape	voldt
Transgenderism	transkøn
Forced abortion	abort
Forced sterilisation	sterilis
Bisexuality	biseks
Transvestism	transvest

Table I. List of topics extracted from the 2002 UN guideline on gender-related persecution that constitute the category. The second column lists the terms that were used in this study to query the Danish case summaries for the topics. The selection of the query words is based on a translation into Danish of the terms used in the UN guideline.

To guarantee reproducibility and to provide the research community with a tool for comparing new methods with the ones evaluated in this study, we provide the code used in this study, as well as the sense-making process on GitHub: https://github.com/KristinKalt/ecscw2022_dk_asylum_analysis.

4.2 Characteristics of the Data

The object of this analysis is the publicly available dataset by the Danish Refugee Appeals Board (<https://fln.dk/praksis>). The data consists of summaries of asylum decisions by the Refugee Appeals Board in Denmark between 2004 and 2021. The Refugee Appeals Board is the second institution to assess applications for asylum in Denmark. Only cases rejected by the Danish Immigration Service, the first instance in the process, are automatically referred to the Refugee Appeals Board.

We conducted the analysis on a subset of appeals cases, namely the cases that are publicly available, for which the selection strategy is unknown. The publication follows an *Executive order on rules of procedure of the Refugee Appeals Board* (Justitsministeriet, Denmark (2016)), which states that the secretariat of the Refugee Appeals Board "regularly updates the board's website www.fln.dk with, among other things, summaries of the board's decisions, the board's background material and other relevant information about the work of the Refugee Appeals Board." (translated by the first author). This guideline is rather broad and does not specify the criteria of cases that are selected to be made public. Therefore the data is not representative of the full set of Danish asylum seeking

cases. From a statistical perspective, we cannot make valid general statements about Danish asylum cases of the scraped data. However, an analysis of the origins of the applied categories provides insights into the archival practices in the Danish asylum system.

Following the data conceptualisation strategy described by Kitchin (2014), we describe the data as follows: The dataset is semi-structured. It consists of an unstructured qualitative part composed of free text summaries. It is supplemented by structured and quantitative metadata comprised of three attributes: decision year, country of origin, and asylum motive. Accordingly, while the attributes *decision year* and *country of origin* have single discrete values, the attribute *asylum motive* can be assigned multiple categories. These types of overlapping categories are also referred to as *tags* in data science.

The categories of the attribute *asylum motive* are collectively exhaustive in the sense that each case fits into at least one given category. To make this system work, residual categories such as *other* are often needed to describe cases that do not fit in the carefully designed boxes of the classification scheme (Bowker and Star (1999)). The Danish dataset utilises residual categories as well, such as *other gender-related persecution* or *other conflicts with authorities*.

As the data points are summaries of asylum cases, the data is derived from other data traces of the process of an asylum application in Denmark. The authors are using this data as secondary data, which according to Kitchin (2014) is data made available for reuse by someone other than the people who generated it and for a different purpose. The data is anonymised by substituting names with letters or a broad descriptive noun (e.g. boyfriend, agent) and is therefore considered attribute data, which represents an aspect of the phenomenon of asylum applications in Denmark, but is not uniquely identifiable (Kitchin (2014)).

5 Findings

	Cases	% of total cases
Gender-related persecution	303	3.3
LGBT	211	2.3
Sexual conditions	366	4.0
Marital conditions	474	5.2
Other gender-related persecution	175	1.9

Table II. This table illustrates the distribution of case summaries among the empirical gender-related categories in the dataset of the Danish Refugee Appeals Board.

In what follows, we illustrate how we make inferences about asylum decision-making and archiving practices. Of the 9,075 extracted cases, 1,876 are tagged with at least one of the five categories that are the subject of this analysis. 20% of the total cases have an asylum motive that falls under a category related to gender (see

Table II). An overall look at the distribution of case summaries among the empirical gender-related categories in the case summaries demonstrates that gender is indeed an asylum motive that we need to take into consideration as we work to understand asylum decision-making practices. Our findings show that 1) gender in asylum case data is a co-dependent category, and 2) that some topics related to gender are omitted in the asylum motive categories, while others are articulated.

5.1 Gender-related Persecution as Co-dependent Category

A first observation and finding of our data analysis is that gender appears to be a co-dependent category. We found that none of the empirical categories on gender-related persecution are applied as a standalone motive for asylum in the Danish data. They always occur together with at least one other category, such as e.g. *religious conditions*. We see in Table III how gender-related categories and the other asylum motives are interconnected. The table illustrates the categories that have been applied together in individual cases. The cells with darker shading highlight categories that are tagged together the most. We find that especially the categories of *ethnic*, *political* and *religious conditions*, as well as the more general categories of *private law matters* and *general conditions* occur together with the categories of gender-related persecution.

	Gender-related persecution	LGBT	Sexual conditions	Marital conditions	Other gender-related persecution	Total
Private law matters	116	37	102	171	39	465
Agents of Persecution	79	14	71	89	63	316
General conditions	53	15	46	67	42	223
Religious conditions	20	21	51	57	4	153
Political conditions	32	17	35	29	6	119
Abuse	21	12	41	21	4	99
Ethnic conditions	29	4	23	17	7	80
Harassment	26	13	19	12	1	71
Criminal conditions	10	3	15	7	2	37
Private law conflict	6	1	7	20	3	37
Military conditions	5	2	9	6	1	23
Nationality	4	4	2	10	1	21
Torture	2	3	9	6	1	21
Departure conditions	6	0	5	5	1	17
Connection to opposition groups	0	1	4	6	4	15
Other conflicts with authorities	0	1	4	6	0	11
Detention	5	0	4	1	0	10
Disproportionate punishment	1	0	4	4	0	9
Recipient conditions	0	3	2	1	1	7
Family relations	2	0	1	2	0	5
Health conditions	0	0	0	2	0	2
Asylum procedure	0	0	0	1	0	1
Families with children	0	0	0	0	0	0
Changed conditions	0	0	0	0	1	1

Table III. This table shows the co-occurrence matrix of the five empirical gender-related categories (columns) and how often cases were tagged together with the other 24 asylum motives (rows) in the public dataset of the Danish Refugee Appeals Board.

This confirms the general legal understanding that there is a category hierarchy in refugee law, as exemplified by the UNHCR guidelines, on how states may apply the category of gender-related persecution. It means that asylum is granted (or not), by subsuming gender aspects into the formal set of Convention grounds (race, religion, nationality, membership of a particular social group, political opinion). Gender-related persecution does not constitute a valid claim for asylum on its own. Rather, it is used in connection with other categories to establish a sufficient asylum motive.

5.2 Omitted and Articulated Gender-related Topics in Asylum Motives

A second observation and finding of our analysis of the origins of the empirical categories associated with the formally defined category of gender-related persecution is, that some topics are made visible, while others are omitted in the process of archiving. Table IV shows the distribution of the 11 topics (rows) among the five empirical categories (columns) of the Danish asylum cases. The percentages illustrate how each empirical gender-related category is composed of several topics. Some categories, such as *LGBT* persecution, are more homogeneous than others, in the sense that they are mainly composed of a few topics that contain a high share of cases in that category. The empirical topic of gender-related persecution on the other hand is broader in scope and more heterogeneous, since it is composed of many topics that all have a comparably low share of the cases in that category.

	Gender-related persecution	LGBT	Sexual conditions	Marital conditions	Other gender-related persecution	Total
Homosexuality	9%	89%	12%	1%	1%	269
Rape	17%	12%	21%	30%	19%	262
Female genital mutilation	8%	0%	2%	17%	65%	177
Forced marriage	7%	0%	5%	41%	14%	143
Human trafficking	6%	1%	1%	1%	11%	48
Forced abortion	1%	1%	3%	10%	2%	41
Bisexuality	0%	15%	1%	0%	0%	39
Forced prostitution	4%	2%	2%	1%	5%	35
Transgenderism	0%	2%	0%	1%	0%	6
Forced sterilisation	1%	0%	0%	0%	1%	6
Transvestism	0%	0%	0%	0%	0%	0
Total	303	211	366	474	175	

Table IV. This table shows the 11 topics extracted from the UNHCR guideline on gender-related persecution (rows) and the five empirical gender-related categories in the public dataset of the Danish Refugee Appeals Board (columns). It illustrates the prevalence of different topics in the categories. The percentages indicate the proportion of overall cases in the category that are tagged with the topic. One case can be tagged with several topics and some cases are not tagged with any of the extracted topics. Thus, the percentages of the categories (in the columns) don't add up to 100%. The darker the cell shading, the larger the proportion of cases in that particular category that mention a particular topic.

Taking an archival perspective, and attending to power relationships, we analysed which topics are omitted in the case data and which topics are articulated and thus made visible.

As seen in Table IV, some topics are dispersed over several categories, such as *human trafficking*, *rape*, or *forced prostitution*. These topics are omitted in the database structure. If the database is accessed through one of the empirical categories, only a partial picture of all cases related to that topic is provided.

The topic of female genital mutilation is predominantly found in the residual category of *other gender-related persecution*. This topic is also omitted. Classification schemes often depend on residual categories to render themselves complete. Residual categories cover the cases that do not fit into *pure* categories (Møller and Bjørn (2011) following Bowker and Star (1999)). One thus expects a wide scope of topics to be included in that category. However, the empirical residual category in the Danish asylum case data, *other gender-related persecution*, is dominated by the topic of female genital mutilation. 65% of the 175 cases in that category cover this topic. In the context of archiving as the act of organising and preserving knowledge, members of a residual category are more invisible than other categories.

Other topics are given prominence by being subsumed under a homogeneous category with a narrow scope. Table IV illustrates that the topics of *homosexuality* and *bisexuality* are articulated in the category of *LGBT*. This category is mainly composed of these two topics.

In the context of an archive, categories take on the role of gatekeepers. They make the data records tractable and serve as the basis for further processing and access. In the context of data science, archival categories turn into features of computational models that aim to predict and inevitably also shape the future. A topic that has been omitted and thus rendered invisible in an archive, has therefore a reduced potential to influence the future when seen in a context of data-driven technologies. Individuals that are part of that category are less likely to be seen and engage with shaping the world. Møller et al. (2021) refer to data shadows as realities that are not captured in tracking data. Here we show how data shadows can also occur in classification schemes.

6 Discussion: Categorisation in Asylum Data and the Context and Practices that Shaped It

Star and Strauss (2004) argue that no category or work practice is inherently visible or invisible but something that we construct. As we see a shift towards casework becoming data-driven (Ammitzbøll Flügge et al. (2021), Saxena et al. (2020)), it is even more critical to recognize the visibility of categories as an arena of power and to deconstruct the origins of data structures (Star and Strauss (2004)). In this sense, the perspective of the archiver (in our case the the Refugee Appeals Board)

is privileged, as they decide which topics are made visible and accessible by being turned into a category and other topics are omitted and therefore difficult to access.

Such omitted topics are characterised by prior CSCW research as data shadows. Møller et al. (2021) establish them as an effective approach for sensemaking of data about a work practice in which stakeholders have differential power. By showing omitted topics (e.g. female genital mutilation) that are not immediately searchable in the database of asylum motives, data shadows appear in our data produced in and through the casework in the area of asylum. We were able to identify these omitted topics by tracing the origins of the five empirical gender-related categories: 1) gender-related persecution, 2) LGBT 3) sexual conditions, 4) marital conditions, and 5) other gender-related persecution (see Table IV). What we learn from the analysis of these empirical gender-related categories is that omitted topics, that do not emerge as a category on their own, are as important to map out as the articulated categories. These are the constituents of residual categories, but also topics dispersed over several categories such as e.g. human trafficking, that enable us to point to specific limitations of a large dataset.

Data shadows are often cast by the categories that are articulated (see Table IV), for example, the topic of homosexuality. Our analysis shows that 89% of the cases in the LGBT category are about *homosexuality*, which thus constitutes a homogeneous category and makes the topic of homosexuality visible. Heterogeneous categories with a wide scope can also cast shadows on the topics they are composed of. The empirical category of gender-related persecution in the Danish dataset casts such a shadow, as it is composed of most of the topics that were subject to this analysis, e.g. rape, human trafficking, homosexuality, and female genital mutilation.

Studying the use of large-scale data in work domains characterised by sensitivity and security concerns such as asylum (Bigo (2014)) is important to ensure they are included in the research on public sector decision-making. We apply NLP to deconstruct gender-related categories as a starting point for engaging with asylum decision-making as a practice. We take an archival perspective to acknowledge power as a factor that shapes the application of gender-related categories. A possible next step is to trace the effects of omitted topics. For example, in regard to female genital mutilation: Is the recognition rate of asylum lower for this particular group or has this topic been omitted because asylum has rarely been granted based on this motive? What does it mean to connect this insight to practice by taking a participatory approach? How can a participation-inspired NLP unfold in a complex, collaborative work domain such as asylum, where researchers act as intermediaries between interests of stakeholders with differential power?

In a design context, *categories* become features of data models that shape caseworker systems. Taking practice as our starting point, CSCW researchers need to find new ways of enabling practitioners to have a reasonable say over the design of these caseworker systems. In our case, that means taking a participatory approach to NLP as suggested by Caselli et al. (2021) and continuing encounters with the stakeholders of the Danish asylum system, such as the NGOs and Danish

national authorities, to further develop the questions we raise in relation to the data. The opportunistic nature of the study in this area with limited access means for example discussing our assumptions around specific categories while being walked through the asylum registration procedure in the Danish reception center.

7 Concluding Remarks

In this paper we set out to investigate how gender-related categories are presented in Danish asylum case summaries and how they can be probed for understanding the practices of asylum decision-making.

We illustrate how data science techniques and NLP can be useful for exploring omitted and articulated topics in asylum case summaries. We uncover data shadows; that is, realities that are not captured by data, such as the topic of female genital mutilation, which is rendered invisible by being organised under the residual category of *other gender-related persecution*. While these omitted topics have a reduced potential to engage in shaping the world in the context of data-driven technologies, being included in official archives is not the only way of being recognized. There are alternative archives that enable communities whose stories haven't been included in official archives to have an impact on the future (Vlachokyriakos et al. (2021)).

We also show how to investigate the empirical reality of a category hierarchy introduced by the historical omission of gender-based persecution as part of the definition in the 1951 Refugee Convention. Despite subsequent state support to include gender as a basis for granting asylum, our analysis shows how category hierarchies are nonetheless routinely reproduced in everyday asylum decision-making. Using a co-occurrence matrix, the analysis shows that in the Danish system, asylum motives related to gender always occur in connection with other categories. As socio-legal research confirms, this has implications for the practice of this domain, specifically for how cases are handled and evaluated by authorities (Spijkerboer (2000)).

This study has some limitations. We only studied a subset of appeal cases, for which the criteria for selection was not known. The NGO we worked with (the Danish Refugee Council) confirmed that not all cases are recorded in the archive. Furthermore, we are cis-gendered scholars with European citizenship and an academic education. That means we are speaking from a position of privilege and engage in an act of knowledge reorganising that is shaped by our standpoint (Harding (2004)). A critical next step for this research is thus to enable the contestation of our findings through participation of stakeholders of the asylum decision-making process.

We invite other researchers to engage with our code (published on GitHub) used for the analysis - as well as the open dataset - as a strategy for knowledge production in this area. Normatively, this study calls on the CSCW community to combine datasets and include data science techniques to study work practices, even if access is limited.

Acknowledgments

We thank Nanna Bonde Thylstrup for the insightful discussion about archival theory and Anna Højberg Høgenhaug for providing her first-hand knowledge into the publication practices of the Refugee Appeals Board. A big thank you to Hubert Zajac for his great advice on the interpretation of the topic analysis. We are also thanking Vasilis Vlachokyriakos, Asbjørn Ammitzbøll Flügge and Trine Rask Nielsen for providing valuable feedback on the overall study. And last but not least, thanks to David Struthers and Colleen Jankovic for copy editing.

References

- Ackerman, M. S., J. Dachtera, V. Pipek, and V. Wulf (2013): 'Sharing Knowledge and Expertise: The CSCW View of Knowledge Management'. *Computer Supported Cooperative Work (CSCW)*, vol. 22, no. 4, pp. 531–573.
- Ackerman, M. S. and T. W. Malone (1990): 'Answer Garden: A Tool for Growing Organizational Memory'. In: *Proceedings of the ACM SIGOIS and IEEE CS TC-OA Conference on Office Information Systems*. New York, NY, USA, pp. 31–39, ACM Press.
- Ammitzbøll Flügge, A., T. Hildebrandt, and N. H. Møller (2021): 'Street-Level Algorithms and AI in Bureaucratic Decision-Making: A Caseworker Perspective'. *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW1.
- Aragon, C., S. Guha, M. Kogan, M. Muller, and G. Neff (2022): *Human-Centered Data Science - An Introduction*. MIT Press.
- Bannon, L. J. and K. Kuutti (1996): 'Shifting perspectives on organizational memory: from storage to active remembering'. In: *Proceedings of HICSS-29: 29th Hawaii International Conference on System Sciences*, Vol. 3. pp. 156–167, IEEE.
- Bechmann, A. (2019): 'Data as Humans: Representation, Accountability, and Equality in Big Data'. *Human Rights in the Age of Platforms*.
- Bigo, D. (2014): 'The (in)securitization practices of the three universes of EU border control: Military/Navy – border guards/police – database analysts'. *Security Dialogue*, vol. 45, no. 3, pp. 209–225.
- Bowker, G. C. (2010): 'The Archive'. *Communication and Critical/Cultural Studies*, vol. 7, no. 2, pp. 212–214.
- Bowker, G. C. and S. L. Star (1999): *Sorting Things Out: Classification and Its Consequences*, Inside Technology. MIT Press.
- Boyd, D. and K. Crawford (2012): 'Critical Questions for Big Data'. *Information, Communication & Society*, vol. 15, no. 5, pp. 662–679.
- Byrne, R. and T. Gammeltoft-Hansen (2020): 'International Refugee Law between Scholarship and Practice'. *International Journal of Refugee Law*, vol. 32, no. 2, pp. 181–199.
- Cakici, B., E. Ruppert, and S. Scheel (2020): 'Peopling Europe through Data Practices: Introduction to the Special Issue'. *Science, Technology, & Human Values*, vol. 45, no. 2, pp. 199–211.

- Caselli, T., R. Cibin, C. Conforti, E. Encinas, and M. Teli (2021): 'Guiding Principles for Participatory Design-inspired Natural Language Processing'. In: *Proceedings of the 1st Workshop on NLP for Positive Impact*. Online, pp. 27–35, Association for Computational Linguistics.
- D'Ignazio, C. and L. F. Klein (2020): *Data Feminism, Strong Ideas*. MIT Press.
- Foucault, M. (1972): *The archaeology of knowledge*. Tavistock Publications.
- Gilliland, A. J. (2017): "'A Matter of Life and Death: A Critical Examination of the Role of Official Records and Archives in Supporting the Agency of the Forcibly Displaced'". *Journal of Critical Library and Information Studies*, vol. Retrieved from <https://escholarship.org/uc/item/5787j3qd>.
- Gitelman, L. (ed.) (2013): *"Raw Data" Is an Oxymoron, Infrastructures*. MIT Press.
- Harding, S. G. (2004): *The Feminist Standpoint Theory Reader: Intellectual and Political Controversies*. Routledge.
- Hathaway, J. C. and M. Foster (2003): 'Membership of a particular social group'. *Int'l J. Refugee L.*, vol. 15, pp. 477.
- Hathaway, J. C. and M. Foster (2014): *The Law of Refugee Status*. Cambridge University Press, 2 edition.
- Irani, L., J. Vertesi, P. Dourish, K. Philip, and R. E. Grinter (2010): 'Postcolonial computing: a lens on design and development'. In: *Conference on Human Factors in Computing Systems - Proceedings*, Vol. 2. pp. 1311–1320, ACM Press.
- Ismail, A. and N. Kumar (2018): 'Engaging Solidarity in Data Collection Practices for Community Health'. *Proceedings of the ACM on Human-Computer Interaction*, vol. 2, pp. 1–24.
- Justitsministeriet, Denmark (2016): *Bekendtgørelse om forretningsorden for Flygtningenævnet*.
- Karusala, N., A. Vishwanath, A. Kumar, A. Mangal, and N. Kumar (2017): 'Care as a Resource in Underserved Learning Environments'. *Proceedings of the ACM on Human-Computer Interaction*, vol. 1.
- Kitchin, R. (2014): *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. SAGE.
- Muller, M., C. T. Wolf, J. Andres, M. Desmond, N. N. Joshi, Z. Ashktorab, A. Sharma, K. Brimijoin, Q. Pan, E. Duesterwald, and C. Dugan (2021): 'Designing Ground Truth and the Social Life of Labels'. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA, ACM Press.
- Møller, N. H. and P. Bjørn (2011): 'Layers in Sorting Practices: Sorting out Patients with Potential Cancer'. *Computer Supported Cooperative Work (CSCW)*, vol. 20, no. 3, pp. 123–153.
- Møller, N. H., G. Neff, J. G. Simonsen, J. C. Villumsen, and P. Bjørn (2021): 'Can Workplace Tracking Ever Empower? Collective Sensemaking for the Responsible Use of Sensor Data at Work'. *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, pp. 1–21.
- Neff, G., A. Tanweer, B. Fiore-Gartland, and L. Osburn (2017): 'Critique and Contribute: A Practice-Based Framework for Improving Critical Data Studies and Data Science'. *Big Data*, vol. 5, no. 2, pp. 85–97.

- Nielsen, T. R. and N. H. Møller (2022): 'Data as a Lens for Understanding what Constitutes Credibility in Asylum Decision-making'. *Proceedings of the ACM on Human-Computer Interaction*, vol. 6, pp. 1–23.
- Pine, K. H. and M. Liboiron (2015): 'The Politics of Measurement and Action'. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. pp. 3147–3156, ACM Press.
- Randall, D., R. Harper, and M. Rouncefield (2007): *Fieldwork for Design: Theory and Practice*. Springer.
- Ring, A. (2014): 'The (W)Hole in the Archive'. *Paragraph*, vol. 37, no. 3, pp. 387–402.
- Saxena, D., K. Badillo-Urquiola, P. J. Wisniewski, and S. Guha (2020): 'A Human-Centered Review of Algorithms Used within the U.S. Child Welfare System'. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA, p. 1–15, ACM Press.
- Seaver, N. (2017): 'Algorithms as culture: Some tactics for the ethnography of algorithmic systems'. *Big Data & Society*, vol. 4, no. 2.
- Spijkerboer, T. (2000): *Gender and Refugee Status*. Ashgate.
- Star, S. L. and A. Strauss (2004): 'Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work'. *Computer Supported Cooperative Work (CSCW)*, vol. 8, pp. 9–30.
- Suchman, L. (1993): 'Do Categories Have Politics? The Language/Action Perspective Reconsidered'. In: *Proceedings of the Third Conference on European Conference on Computer-Supported Cooperative Work*. USA, p. 1–14, Kluwer Academic Publishers.
- Taylor, A., D. Sweeney, V. Vlachokyriakos, L. Grainger, J. Lingel, S. Lindley, J. Lingel, and T. Regan (2015): 'Data-in-Place: Thinking Through the Relations Between Data and Community'. In: *CHI '15, Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. pp. 2863–2872, ACM Press.
- Thylstrup, N. B. (2022): 'The ethics and politics of data sets: deleting traces and encountering remains'. *Media, Culture & Society*.
- Thylstrup, N. B., D. Agostinho, A. Ring, C. D'Ignazio, and K. Veel (eds.) (2021): *Uncertain Archives: Critical Keywords for Big Data*. MIT Press.
- UNHCR (1951): *The Refugee Convention, 1951*. Geneva, Switzerland.
- UNHCR (2002): *Guidelines On International Protection: Gender-Related Persecution within the context of Article 1A(2) of the 1951 Convention and/or its 1967 Protocol relating to the Status of Refugees*. Geneva, Switzerland.
- Vlachokyriakos, V., C. Crivellaro, H. Kouki, C. Giovanopoulos, and P. Olivier (2021): 'Research with a Solidarity Clinic: Design Implications for CSCW Healthcare Service Design'. *Computer Supported Cooperative Work (CSCW)*, vol. 30, pp. 757–783.

Syed, H. A.; Schorch, M.; de Carvalho, A. F. P.; Rutz, P. Pipek, V. (2022): Blending Practices to Facilitate Grounded Design Research: A Praxeological Research Perspective. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Notes, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_n04

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_n04

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

Blending Practices to Facilitate Grounded Design Research: A Praxeological Research Perspective

Hussain Abid Syed, Marén Schorch, Aparecido Fabiano Pinatti de
Carvalho, Philipp Rutz, Volkmar Pipek

University of Siegen, Germany

{hussain.syed, maren.schorch, fabiano.pinatti, philipp.rutz, volkmar.pipek}@uni-siegen.de

Abstract. In this paper, we reflect on the experiences from two Grounded Design (GD) research projects conducted by a multidisciplinary group of researchers between 2019 – 2021 and highlight the methodological foundations and related obstacles for iterative designing. Both projects investigate the phenomena of knowledge sharing and crisis-related learning in business organizations under the GD paradigm, which has been increasingly adopted within the Computer-supported Cooperative Work (CSCW) community. During these projects, the researchers with backgrounds in computer science, business informatics, software engineering, and sociology experienced the need for systematization to transition between the stages of GD. Looking back, we realize that our teams arrived at this systematization by blending the prior knowledge from team members' original educational backgrounds. While blending practices most likely happens intuitively in interdisciplinary projects, as is often the case of the user-centered design initiatives seen in CSCW and Human-Computer Interaction, little can be found on how this usually happens and its implications. In this paper, we respond to this literature gap by discussing how this blending can facilitate the realization of GD projects and lead to a praxeological information science research perspective, which has 'methods appropriation' as key to systematizing abstraction, broader traceability, and flexibility of research methods.

Introduction

Practice-centered traditions existing within sub-fields of Human-centered Computing (HCC), as is the case of Computer-supported Cooperative Work (CSCW) and Human-Computer Interaction (HCI), lie in the fact that a pool of designers or developers does not imagine problems and their respective solutions. Instead, solutions should be grounded in the social aspects of human and organizational behaviors (Kuutti and Bannon, 2014; Wulf et al., 2015). Understanding the contextual needs around a specific problem is implicated in extending the boundaries of human and organizational capabilities by creating new and innovative artifacts. These traditions gave birth to Grounded Design (GD) which advocates a praxeological and pragmatic perspective in the design lifecycle, asserting that the changes in the design must be guided by purposefulness and knowledge of the social practices under investigation (Goldkuhl, 2012; Rohde et al., 2017). The turn to practice and praxeology in GD underlines that information technology (IT) artifacts are adopted and activated for effective use in the organizational context with behavioral purposefulness and, therefore, different (new) social practices emerge as a result of technology adoption and appropriation (Brödner, 2009; Stevens et al., 2010). This practice-centered approach has been receiving increasing attention in the CSCW community and has been used to predicate many and various studies and design initiatives, which have been advancing the state of the art considerably (Rohde et al., 2017; Wulf et al., 2015; Wulf et al., 2018).

The design approaches commonly used in practice-centered computing under GD have a similar set of activities from three to five stages, as will become evident across the paper. GD asserts the iterative and evolutionary nature of practice-centered IT design as a multi-layer intervention in an organization's social practices to better cope with the problem of self-referentiality (Rohde et al., 2017). The self-referential nature of IT design emphasizes that as the prototype is designed to be evaluated and appropriated by and with the users, it also changes the social practices for which the artifact is being created. This leads to frequent changes in user needs and requirements, and therefore, an iterative mechanism must incorporate the changes. GD research ensures this continuity in context and appropriation studies while the design study is subject to alterations. While this view resonates with the user-centered design (UCD) views from HCI and CSCW (Benyon, 2019; Sharp et al., 2019; Harper et al., 2008), it contradicts the design-centered view in information science (IS), computer science (CS), and software engineering (SE) traditions where formal specifications for design can be gathered beforehand and can be reflected in the artifact following a sequence of steps in the design lifecycle (Baskerville et al., 2018; Chandra et al., 2015; W. Kuechler and Vaishnavi, 2012).

The multistage design approaches in the context of GD always own a level of abstraction and, thus rightly so, provide the researchers a wide-open ground to try different techniques within the stages of research. However, the evolutionary nature of GD blurs the boundaries and segregation between the steps of the applied design approach, causing the transition between these stages with a reasonable probability of unsystematic actions causing wastage of research resources. Also, these design approaches can be challenging for researchers with different academic and methodological training transitioning into practice-centered research. It may be due to their lack of in-depth comprehension of diverse, overlapping disciplines that make up the epistemological foundations of practice-centered computing or the shortcomings of the established development lifecycle methods in their parent disciplines. Further obstacles to handle cover, e.g., coping with the incoming knowledge about practices in ongoing research and the lack of mechanisms for systematic transition between stages in the applied design approach. Based on that reflection, we ask: “How can blending practices of researchers from disciplines with dissimilar training in methods facilitate Grounded Design (GD) projects?”

In this paper, we set out to answer this question by presenting experiences from two GD projects implemented respectively through design thinking (Brown, 2008) and Design Case Study (DCS) (Wulf et al., 2015; Wulf et al., 2011). The projects were undertaken by researchers with backgrounds in business informatics (BI), CS, and sociology. By carefully reflecting on our experiences, we argue that the methodological flexibility and the evolutionary nature of GD approaches demand a systematization in methods selections, which is heavily influenced by the prior methodological training of the people working on the project. Our conclusions suggest that the background knowledge of the researchers guided the ‘method appropriation’ to systematize the transition between the context study and design study stages of the applied design approaches. Furthermore, we state that systematization led to traceability between project stages and results, which is critical for the success of GD projects. We offer, therefore, a methodological contribution to the field, which can support researchers and practitioners engaging in GD projects to avoid potential hurdles as they move forward with their initiatives.

The forthcoming sections include related work about methods in practice-centered design traditions and other dissimilar research and design fields, the explanation of the research contexts and experiences in our two GD research projects, and building on that foundation, the discursive contribution of method appropriation as implications for GD research, followed by a conclusion.

About methods: Related work

Since our contribution refers to blending practices from different fields and traditions, we find it relevant to provide an overview and contextualization of the methods and approaches we refer to as we construct our contribution. We start with a brief historical account of the approaches used in HCI and CSCW across the years and then address some approaches from CS and IS, highlighting some methods from SE on which we drew during the cases that we present.

Design approaches in HCI and CSCW

Since its inception, HCI has put considerable effort into devising approaches to address the design of computer technologies from a human-centric perspective. An essential characteristic of these approaches is the use of methods from disciplines like psychology and sociology, where the roots of the field can be found. CSCW, which emerged from the change from a human factor to a human actor perspective to HCC (Bannon, 1995), shared many of these approaches, especially concerning the use of qualitative methods to understand and react to the socio-technical aspects of the design. In turn, GD has been proposed as a paradigm for HCC computing, providing a set of ontological and epistemological assumptions to guide design research in learning about the users, their world, and their needs (Rohde et al., 2017). It adds the evolutionary stance to HCI and CSCW design research to be conducted in an agile manner incorporating the emergent changes in social practices that were a result of appropriating and using IT artifacts, hence organizing the practice and technology development as an integrated process (Wulf et al., 2011; Rüller et al., 2021).

While GD has been proposed as a paradigm for HCI and CSCW research and practice, DCS have been proposed as a framework for GD projects (Wulf et al., 2011; Rohde et al., 2017; Wulf et al., 2015). The framework is organized in three steps: in a *pre- or contextual study*, predicated on qualitative methods like observations and interviews, that provides researchers with a first-hand opportunity to learn and better understand users' behaviors and interactions within a particular context. Quantitative methods can also be employed in this phase in a mixed methods approach, though they are not that common. The results of the data analysis form the basis for design implications and, later, for prototyping. The *design* phases usually follow a participatory design (PD) approach, which entails the active engagement of the users in the conceptualization of the design ideas using various methods, like (collaborative) prototyping at different fidelity levels, co-development, card games, contextual enquire, to name but a few (Muller and Kuhn, 1993).

In the *appropriation*, a stable version of the prototype is tested with the users in their natural organizational settings for the appropriation and usage of the prototype

and its transformative impact on the focused everyday practices and the design of the prototype itself. The collected data during this phase is again analyzed and is transformed into further design implications for the improvement of the prototype. Although the DCS phases initially occur in sequence, as one step depends on the results of previous stages, these phases partly coexist across the GD initiatives.

While DCS has been traditionally associated with GD, it is not the only approach that can be used to implement GD initiatives. Any CSCW and HCI approaches that allow for the understanding of practices and the use of such understanding for the design and assessment of computer technologies conform with the premises of GD. An example of it is *design thinking* (Brown, 2008). In fact, the five phases of design thinking – *empathize*, *define*, *ideate*, *prototype*, and *test* – can be easily mapped to the first two phases of DCS: *pre-study*: *empathize* and *define*; *design*: *ideate*, *prototype*, and *test*. Nevertheless, the appropriation phase falls out of the scope of this framework – although some would argue that the test phase can also be implemented as an appropriation study. DCS can also be easily mapped in other design methodologies, such as the ISO 13407 UCD process and associated interaction design processes that have been proposed over the years consisting of iterative process for evolutionary development (Benyon, 2019; Sharp et al., 2019). UCD process provides the baseline template for iterative development to incorporate the feedback and changes in requirements from the users, pointing out the stages in UCD being affected by these alterations but does not elicit a structured way to achieve this evolutionary process. This hinders the practice of UCD as an end-to-end design methodology due to restricted scalability and high level of abstraction for practitioners (Mao et al., 2005).

All these methods and frameworks have been beneficial – at least in theory – to help HCI and CSCW professionals to engage with the many issues concerning the field seriously. Nevertheless, as we argue across this paper, there is a fair amount of flexibility and abstraction in these instruments that, although positive in some cases, can be negative in other cases, especially regarding less experienced researchers, who can have difficulties instantiating the appropriate methods in their projects.

Prominent design approaches in other design traditions

Design methods and approaches have evolved considerably in CS/SE since Dijkstra's idea of top-down *structured design* disintegrated the application's functionality iteratively into multiple layers of sub-functions until the hierarchy of subroutines rather than an assembly of computer instructions is achieved (Dijkstra, 1968). As an alternative, Hoare (1971) presented a four-stage method of *formal specifications* focusing on translating the requirements into the semantic logic detecting the defects before software development. These efforts were parallely met with the formulation of software development life cycle methodologies with a

project management perspective, initially by Royce (1970), who introduced the *waterfall model* with stages like requirements analysis, design, development, testing, and maintenance. Customer satisfaction as a factor in the design life cycle was introduced by Basil and Turner (1975) in an *iterative development method* that solicited developing one component at a time and improving the design by incorporating early user feedback. The concept of *rapid prototyping* by Luqi (1989) further elevated the notion of iterative development, which included the customer requirements into a prototype at first, that can be evaluated and refined, later to be developed as the final product. The success of rapid prototyping techniques and the proliferation of Internet-bound applications encouraged the development of *agile methods* following the stages of rapid prototyping in parallel and overlapping design iterations with short-term and version-bound product releases (MacCormack et al., 2001).

In this respect, such design and development methodologies were influenced by project management perspectives of fast-tracking an error-free product to market. The users were considered, but the social practices of users were not given the center of attention as it is done in practice-centered computing. On the other hand, disciplines like IS and Design Science (DS) pursued the ideology of '*learning through the act of building*' that emphasizes creating a design theory through the process of developing and testing IS artifacts which is inextricably bound to the testing and refinement of its kernel theory (Gregor, 2006; W. Kuechler et al., 2005). Keeping in view the traditions of IS research, Kuechler and Vaishnavi extended the activity framework of DS research by incorporating a cyclic knowledge flow between the steps of their design method to facilitate the theory creation by addressing the constraints in the design process (B. Kuechler and Vaishnavi, 2008; W. Kuechler and Vaishnavi, 2012). The method consisted of five steps: *awareness of the problem, suggestion, development, evaluation, and conclusion*. The *awareness of the problem* construes conceptualizing the problem space, the goal, and the scope of the DS research project. The knowledge gained informs the *suggestions* for the *prototype*, which is to be *tested* and *evaluated* with the users. The *conclusion* step finds the results to support the next design iteration.

Feine et al. (2020) elaborated the *suggestion* step of W. Kuechler and Vaishnavi's method by conducting a formalized requirements engineering for a systematic transition into the development stage. The illustration of this adaptation can be seen in Figure 1. They systematized the suggestion step by reviewing their empirical data to conceptualize *crude requirements* for the artifact design. Based on the requirements, they drew *design principles* following the definition of Chandra et al. (2015), introducing "statements that prescribe what and how to build an artifact to achieve a predefined design goal" (p. 4040). Finally, the linkage of requirements and design principles is further extended into concrete *design features* as per Seidel's notion of substantiating technical specifications into concrete design elements for the prototype (Seidel et al., 2018).

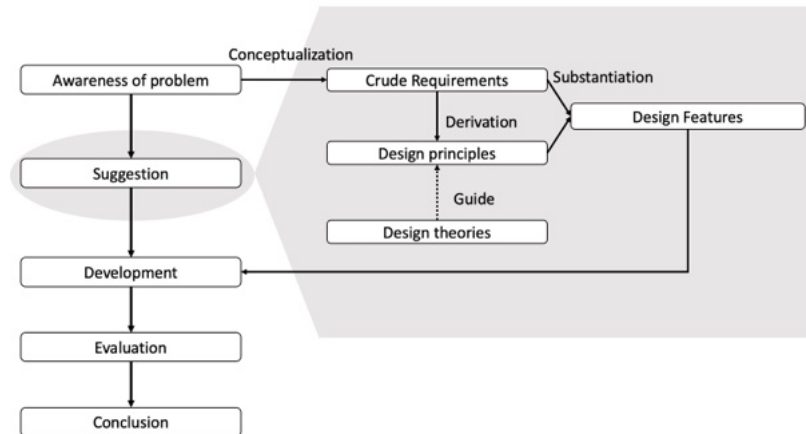


Figure 1. Adaptation of W. Kuechler and Vaishnavi's method for systematic transition between problem and development (Feine et. al., 2020).

The task to systematize the knowledge flow between problem and design also resulted in formalized methods in BI by Balzert. According to Balzert (2010), the requirements analysis is based on a two-stage procedure in which first the crude requirements and then the functional specifications are created. The crude requirements are abstract conceptualizations of evidence-based justifications for design based on the problem statement. They hold the visions, goals, design concepts, and framework conditions to be developed after refinement through operationalization. The refinement of crude requirements substantiates the specifications, which are the concrete design features for development. In it, the specific functional and non-functional requirements, such as quality requirements, are formulated based on the crude requirements. The illustration of the method is shown in Figure 2.

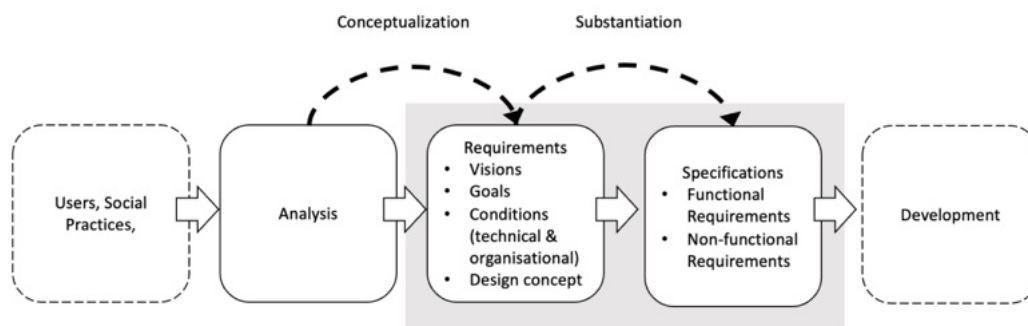


Figure 2. Balzert's method of requirements elicitation from crude requirements to specifications for design and development.

When noting these crude requirements and functional specifications, Balzert's method follows a fixed notation scheme:

/Letter Number/ Description of the requirement

The *Letter* refers to the type of requirement. At the same time, a unique *Number* is assigned consecutively for each new requirement (e.g., **/F10/ *The application should always be accessible via the Internet or the intranet***). This pseudonym scheme systematizes transitions between development stages by linking the specifications to crude requirements and further back to problem statements.

The ‘systematization of knowledge transition’ as a bridge between problem and development stages in the methods of Balzert and the adaptation of Kuechler and Vaishnavi serve as a guiding ideology for ‘method appropriation’ in Project 1 and Project 2, respectively, which are explained in the section below. These methods appropriation is with noticeable differences as per GD research that the awareness of the problem in our case is generated by an inductive ethnography of social practices over an extended period. The underlying design approach is iterative, and the research is evolutionary.

Methods Appropriation: A Projects-based Exploration

Our contribution is based on two GD projects (labeled Project 1 and Project 2) undertaken by a group of researchers and designers with a multidisciplinary background. Both projects were parts of a larger interdisciplinary research project, situated in an IS department at a German university. The larger project dealt with determining the conditions for maintaining continuity and supporting resilience in small and medium-sized enterprises (SMEs) and designing IT solutions to support the SMEs in enhancing their resilience against crises. The sub-projects were undertaken by teams of two and three researchers from the research group while employing the design thinking approach and the DCS framework, respectively, with a common GD research mindset. During these projects and as problems emerged, a series of research methods have been appropriated, i.e., they have been adapted and fit to the working practices following integration in the process (Dourish, 2003). An abstraction of the design approaches, method of data collection, analysis, prototyping, and evaluation is shown in Table 1.

We describe the two projects’ outlines and the applied methods and practices in the following. This builds the foundation for our methodological implications. As this paper does not focus on the domain-specific thematic and design implications within the two projects but presents the methodological implications for using, adapting, and appropriating design approaches under GD, the explicit domain-specific findings about the designed artifacts (solutions) are mentioned concisely, enough to generate understanding about the research contexts.

Table 1. Overview of methods applied and outcomes in the two projects

Project 1: Platform for crisis knowledge sharing			
Design thinking approach			
	Empathize and ideate	Design and prototype	Evaluation
Methods	<ul style="list-style-type: none"> - Semi-structured interviews - Observations - Thematic analysis - Brainstorming 	<ul style="list-style-type: none"> - Evolutionary prototyping - Design workshops 	<ul style="list-style-type: none"> - Pluralistic walkthrough - Thinking aloud method
Outcomes	<ul style="list-style-type: none"> - Interview transcripts - Observation protocols - Codes and themes - Mind maps - Design implications 	<ul style="list-style-type: none"> - Low fidelity prototype - High fidelity prototype 	<ul style="list-style-type: none"> - Implications for appropriation
Project 2: Application for crisis learning			
Design case study framework (DCS)			
	Pre-study	Design	Appropriation
Methods	<ul style="list-style-type: none"> - Open and semi-structured interviews - Observations - Thematic analysis 	<ul style="list-style-type: none"> - Evolutionary prototyping - Pluralistic walkthrough 	<ul style="list-style-type: none"> - Pluralistic walkthrough - Follow-up interviews
Outcomes	<ul style="list-style-type: none"> - Interview transcripts - Observation protocols - Codes and themes - Design implications 	<ul style="list-style-type: none"> - Low fidelity prototype - High fidelity prototype 	<ul style="list-style-type: none"> - Implications for appropriation

Project 1: Platform for crisis knowledge sharing

This first sub-project took place for a year between 2020-2021 and was carried out by a team of two researchers (Ph.D., Masters) having CS and BI backgrounds, respectively. They explored the phenomenon of knowledge sharing for crisis preparedness in a metal structures construction and maintenance company. Here, design thinking (Brown, 2008) was chosen as the overall approach to understand the problem space in the company. The approach was employed considering the premises of GD and its praxeological orientation, focusing on improving the social practices and corpus of situated design cases subject to meta-analysis (Rohde et al., 2017). In particular, the stages of design thinking were done in a non-linear and iterative fashion to support the evolutionary research based on this project.

Research context and empirical study

The pragmatic nature of GD requires that practice is grounded in background knowledge that is not entirely explicit and relates to emotional and motivational elements (Kuutti and Bannon, 2014). Therefore, we drew on an inductive approach with reference to ethnography and, more precisely, observations and semi-structured interviews with stakeholders of this context (Silverman 2011; Flick 2018; Strauss and Corbin 1998; McDonald 2005). The overall project objectives (topics of crisis management and knowledge-sharing) worked as guidelines for the interviews and in some abstract way as background for the observations.

One of the researchers went for a full-time four-month internship in the company to carry out observation through shadowing (Quinlan, 2008). The researchers collaboratively conducted six semi-structured interviews with employees in the company on topics like crisis communication, crisis knowledge accumulation, and sharing. The internship was done within the company to analyze and participate in the design process of the knowledge-sharing system within the action context of users (McDonald, 2005). The focal point of the internship was more general from a knowledge sharing point of view, but we agreed with the management to use knowledge-sharing for crisis as a use-case for the whole internship. During the internship, the researcher shadowed the knowledge workers in the company, including the document control manager, senior trainer, and security manager. The internee also kept fieldnotes observing the employees in action and recording the formal and informal exchanges between different employees in the company concerning knowledge-sharing and documentation (David Silverman, 2011; Quinlan, 2008). The interview data were transcribed, and the whole data set (including the field notes) was analyzed using inductive thematic analysis (Braun and Clarke, 2006) with the software MAXQDA[®] (Kuckartz and Rädiker, 2019). The authors collaboratively performed open coding to find initial codes and themes which were analyzed further for specific interaction patterns and socio-technical factors impacting the individual, team-level, and organizational practices.

The analysis of knowledge-sharing about crises was crucial to understanding the users' real problem space and its intricacies. But despite being very elaborate, the findings of the analysis left an expansive room for interpretation. The themes did offer perspectives and implications for ideation and design. Still, there was a possibility of several design features for prototyping, all grounded in the practice context of the users. The research team and the partners from the company decided upon the key design elements for the first prototype. Furthermore, extending the design functionality of the prototype iteratively afterwards by utilizing the knowledge from analysis and appropriation.

Systematization for iterative research

The evolutionary nature of the prototype prompted the need to systematize the knowledge transition between analysis and prototype design. This systematization was necessary due to the openness of the extracted knowledge from field research and the interpretative margin the analysis provides for the design implications. The motivation which led to using a systematic approach between empathize/ideation and design/prototype phases of the design thinking approach was to be able to forward and backtrack the design decisions to the design implications and empirical knowledge to facilitate the evolution of the prototype during later appropriation cycles as per GD research.

During brainstorming for ideas to achieve systematization, the *prior background knowledge* of the researchers from the CS and BI fields guided the decision to apply the requirement elicitation method from Balzert (2010), which is suitable for small projects and is also tested in practice (Hatterscheid and Schluter, 2018; Lohrmann and Reichert, 2016; Majchrzak et al., 2018). The method proposes a two-stage model to extract formal and design-specific requirements from abstract and interpretative knowledge. The thematic analysis of the collected empirical data resulted in abstract and non-systematic requirements for the prototype design in themes, design implications, and ideated mind-maps. Balzert's method was used to generate systematic functional and non-functional requirements and framework conditions for the prototype, which can be evolved in the iterative process (see the method in the related work section).

Initially, the crude requirements were generated by conceptualizing the design implications and ideas into the vision of certain design concepts, goals to be met with design decisions, and later substantiating the concrete specifications with functional, non-functional, and quality requirements. The requirements elicitation provided means to backtrack the design decisions to the knowledge generated during empathizing and ideation stages of design thinking and to make informed decisions for design and prototyping. The illustration of blending design thinking with Balzert's method is shown in Figure 3.

This blending of methods seemed essential as the evaluation of the prototype influenced the extension of the prototype further towards system development following the iterations. For example, a consensus between the researchers and the tentative users inspired an initial design of a wiki-based knowledge-sharing platform but with the possibility to adapt, scrutinize, customize, and revert the design features in the prototype as per users' feedback. The need for systematization became evident while incorporating the feedback after the first evaluation cycle, which altered the initial design decisions of having an open, collaborative space for knowledge creation, management, and sharing, to a mechanism with version control and role-based knowledge creation and management. The systematization offered by method appropriation allowed us to address the challenge of incorporating the understanding of altering user needs and

practices (self-referentiality) as the users were also defining new practices and changing the old ones while testing the design iteratively. It allowed us to backtrack the design decisions into analysis and extend the design implications with respect to interconnected components of design.

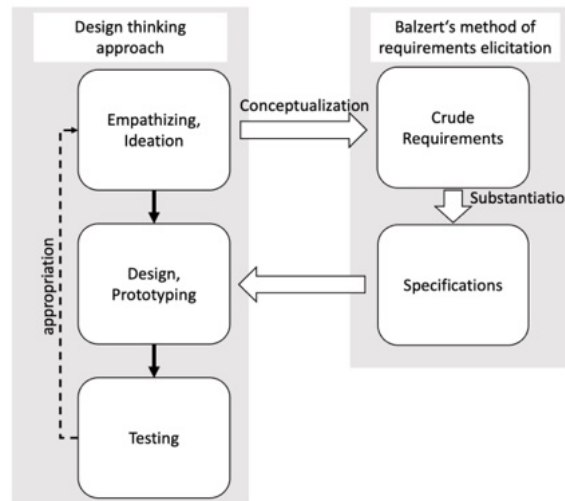


Figure 3. Blending Balzert's method of requirements elicitation within stages of design thinking.

This blending process within the design thinking frame not just permitted to systematize the extraction of formal requirements for a working prototype but also invoked more clarity in the minds of researchers with the confidence that how certain design decisions were supported by empirical evidence and the user needs to be originated through field research. This inductive method guided the researchers through the openness between the design implication or findings of empirical research and the prototyping stage of the research but also turned out to be a more inclusive experience as it aided in bridging the gap between the researcher's prior knowledge from their respective disciplines and newly learned methods in practice-centered design research.

Project 2: Application for crisis learning

The second sub-project focused on crisis-related learning and training processes in business organizations, especially small and medium-sized enterprises. It was undertaken between mid-2019 till mid-2021 and involved a research team of three researchers (Postdoc and Ph.D. candidates) with backgrounds in sociology and CS. The epistemological research interest in this sub-project was an ethnographic approach, covering different qualitative methods for two years. It was realized with the DCS as the principal frame of reference (Wulf et al., 2015). Besides the research interest in gaining insight perspectives about crises experiences and related learning and training practices from the companies, a second goal was to design with the

companies to enhance resilience by improving bottom-up coordination between employees and the management.

Research context and empirical study

Following the pre-study stage of DCS, the researchers started with data collection initially via open and later semi-structured interviews in different local small and medium-sized companies. The interviews were done with employees from the executive to labor positions in the companies. They covered the topics of crisis experiences and management, preparedness, and learning and training practices. In 2019, the interviews were carried out on-site in an open interview format with a pair of researchers from the team (Flick, 2018) and as a contextual inquiry (Holtzblatt and Jones, 2017), i.e., they were also accompanied by tours in the company, which often resulted in observation sessions and informal, spontaneous interactions with the company's employees. In the first nine months, three companies were part of the study, with regular sessions of participatory observations (Silverman, 2011; Flick, 2018; Emerson et al., 2011).

This process became frozen in early 2020 when the Covid-19 pandemic started to evolve. All the companies were suddenly confronted with a real crisis and had to prioritize their crisis management; many had to shut down for months. The several stages of the pandemic with various measures prevented the research team from realizing the planned long-term observations. Like many other researchers, the methodological tool kit was adapted (Self, 2021), and the interviews were done entirely in digital format, using communication tools like Zoom® or Skype®. In addition to the three companies from the pre-pandemic research phase, nine companies were interviewed (semi-structured interviews) over the course of one and a half years (March 2020 – August 2021).

Parallel to the data collection (mid-2019), the researchers started analyzing the transcribed data material by means of thematic analysis (Braun and Clarke, 2006) for fine-grained design implications and broader themes with MAXQDA® (Kuckartz and Rädiker, 2019). The analysis followed an inductive method (creating codes openly from the data, Strauss and Corbin, 1998) and later, with a further developed code scheme, an interplay between inductive and deductive coding. The researchers analyzed the data material in regular data analysis meetings to remove bias and gain inter-subjectivity. After halfway through the project, when the analysis results accumulated meaningful descriptions of the crisis-related learning and training phenomena, the researchers began the prototyping phase of the DCS. The initial obstacle was to translate the thematic understanding of the problem into design implications. Secondly, due to the ongoing research (following GD and DCS guidelines), an evolutionary method was required which could accommodate the upcoming thematic and design guidelines.

Systematization for iterative research

After a series of deliberation and discussion cycles based on *prior knowledge* about methods in DS and CS, keeping in view the application of the methods in successful projects (Chandra et al., 2015; Feine et al., 2020; Seidel et al., 2018), the requirements elicitation of Kuechler and Vaishnavi's adaptation of design science research approach (W. Kuechler and Vaishnavi, 2012) was chosen. The agile nature of the method and the backtracking mechanism it offered from design features to design principles and a further step backward to the requirements in the empirical evidence were the motivating factors behind choosing the method as a transitional approach between pre-study and prototyping. This method's forward and backtracking mechanism supplied means to upscale the prototypical mock-ups during a long-term qualitative empirical study and the DCS appropriation loops. It was necessary to support prototype appropriation with incoming design contexts simultaneously from the data analysis sessions and the evaluation with users. The illustration of blended DCS with Kuechler and Vaishnavi's approach is shown in Figure 4.

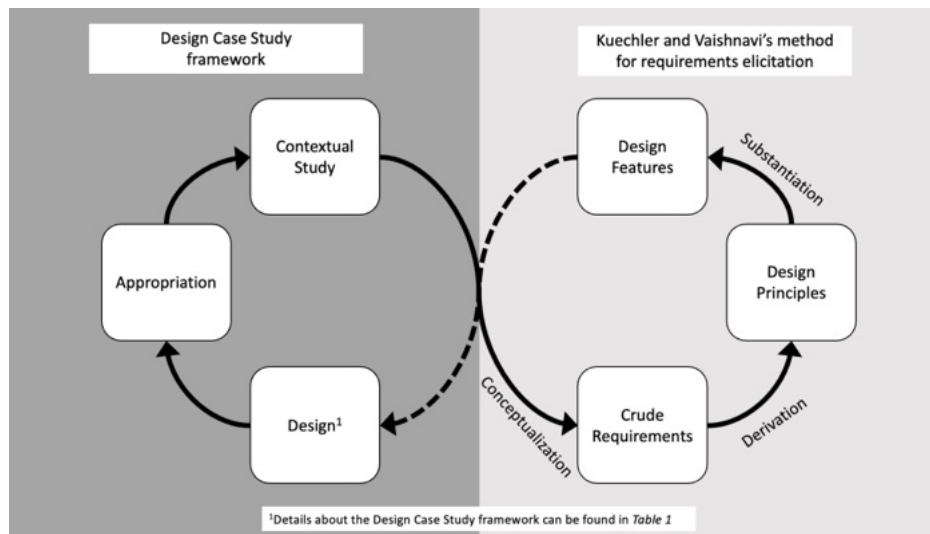


Figure 4. Blending adaptation of Kuechler and Vaishnavi's for requirements elicitation within stages of DCS.

For example, the initial findings from the analysis pointed out the elitist culture in the organizations, where the opportunity for crisis-related learning was designated for a selected few who were deemed responsible for crisis management. This gap instigated designing a platform with easy access to knowledge about crisis response which ensures equal opportunity for employees at all levels in the company. However, the first evaluation study resulted in the need to incorporate means for self-learning to inspire flexibility for crisis-related learning and a feedback mechanism between management and employees to address learners' inquiries about crisis response. These challenges were addressed by extending the

functionality of the initial prototype with design features like chatbot as a learning assistant and inquiry generation for expert-in-the-loop.

The blending of Kuechler and Vaishnavi's approach for a systemized and traceable transition between design implications and the evolutionary prototype eased the realization of interdependencies between existing and new design elements, hence the development of a fully functional high-fidelity prototype that was not just realizing a platform for mutual awareness but also a medium for communication and learning for crises preparedness. The evolution of this prototype between different fidelities and the changing requirements (with continued empirical work and parallel evaluation) expands the horizon of GD research, presenting a compelling case for sustainable development in ongoing research projects and underlines the importance of a traceable and systematic process of design.

Methodological implications for design research: lessons and experiences

“Practice-orientation is a labor-intensive, risky, and long-term research approach” (Wulf et al., 2011, p. 510). The GD paradigm highlights that the underlying problems of existing social practices are not obvious and often need to be shown by extensive observation and analysis (Rohde et al., 2017). It encourages the discourse of cross-sectional investigation across design cases, looking for patterns and similarities to ensure self-referentiality of social practices in long-term and evolutionary IS projects which can support appropriation in several contexts. The iterative nature of design approaches with GD mentality allows design and appropriation activities to intervene in a self-referential circle, meaning that technology and practice can evolve together. It smudges the boundaries between the salient stages of the applied design approach because the incoming knowledge from analysis of ongoing context and appropriation studies needs to be incorporated into the design recursively.

Without a clear transition between the salient stages of the applied design approach and apparatus to forward and backtrack design features and implications of the prototype into analysis and vice versa, leads to unsystematic evolution of the design, which is not self-referential to the changing social practices (Rohde et al., 2017). The first iteration in such GD projects is crucial because it lays the foundation of the evolutionary mechanism in the project for the upcoming iterations where constantly increasing corpora of knowledge about users' social practices need to be reflected in the design. This can be enabled by ‘method appropriation’ right from the beginning in GD projects. This systematization process can only be generalized as a research practice if the group involved in the GD project identifies the potential of going beyond the guidelines of design approaches and capitalizes

on the diverse epistemic palette of the researchers (in the group) as means to navigate through the long-term research objectives, the evolutionary spiral of the design process and inter-projects transferability of findings. We have achieved method appropriation by using our prior knowledge/training about methods/approaches to systematize the transition, as illustrated by the project experiences in the last section. This systematization and blending practices reflect three methodological implications for CSCW research to fully harness the potential of GD research.

These implications add to the conceptualization and realization of GD research promising structured evolution and propagation of knowledge about social practices. The first implication recognizes the flexibility of GD research with the allowance of scholarship from many interrelated disciplines as means to blend researchers' prior knowledge and training. The second implication addresses the abstraction of design approaches used in the GD projects emphasizing the need to systematize knowledge transition between the stages of design approaches to accommodate knowledge reusability and cooperation among stakeholders. Lastly, the third implication illustrates the traceability in GD projects which is also a percussion of systematized knowledge transition addressing the problem of self-referentiality in GD projects.

Flexibility to blend prior knowledge of researchers

Researchers from various fields join CSCW and HCI disciplines for doctoral and post-doctoral research. The researchers are primarily trained in the methods from their respective studies or often learn them in previous jobs (Mydin and Surat, 2021). Adopting the norms and practices of the new discipline is considered crucial for research to progress and publish the work done in research projects effectively. In this contribution, we assert that prior training and experience of the researchers come in handy while learning new methods and doing actual research. This can lead to the adoption and adaption of the methods in HCI and CSCW with a personalized flare in a research framework that can be agile and self-supporting. The examples of method appropriation for design approaches by blending the prior methodological knowledge and training of researchers in this research is an epitome of our stance. This kind of inter-method marriage or focused blending of methods can make research projects flexible and inclusive for interdisciplinary research. Likewise, such flexibility demands and builds on *reflexivity*, emphasizing the role of the researcher – who they are, what they do, and how they shape their own practice, which is an important aspect for the successful conduct of evolutionary research and iterative design processes (Mauthner et al., 2003; Frost 2016). This is entirely in accordance with the GD research initiative of contextualized inter-project coordination and learning for social practices (Rohde et al., 2017). Promoting flexibility in research projects by providing creative space for

researchers to bring along their experiences and knowledge is also in line with building affordances for research through cognition and actualization of prior research practices (Bernhard et al., 2013). As the researchers go on appropriating methods with prior knowledge and experience, they associate themselves with the research on a deeper level, knowing that they can control the evolution and iterations of a long-term process.

Systematizing abstraction in design approaches

The two design approaches in the research projects mentioned above point towards a certain level of abstraction in their organization to provide a guideline for research (Gaver, 2012; Hoök and Lowgren, 2012). Abstraction is a necessity to leave room for adaptation and out-of-the-box thinking. However, the subjective nature of ethnographic research applied to understand social phenomena and practices contextualizes the abstraction. This contextualization is highly reliant on the type of research methods and researchers' perspectives of the phenomenon under investigation (Frost, 2016; Frost et al., 2010; Flick, 2018). The use of well-structured concept-driven methods to communicate the design constructs through the design of artifacts can systematize abstraction (Stolterman and Wiberg, 2010). Then again, the design propositions generated in the ongoing analysis through iterations can inform different and even conflicting design decisions. This subjective decision-making process on behalf of researchers often gets exploited due to the openness of the applied research methods. It can lead to DS projects often not delivering sufficient benefits and newly developed information systems underperforming the expectations (Rohde et al., 2017). GD's evolutionary nature aims to transcend across various iterations of a project and extend the evolving corpora of social practices learned as a part of project scholarship, reflecting the need for systematization.

Systematizing abstraction does not propagate the project management mentality for maximum monetization as in business and software engineering traditions but to mechanize knowledge acquisition and incorporation. The involvement of users in the research projects during pre-study, design decisions, evaluation, and appropriation loops in long-term projects require coordination with the users using different deliverables from various stages of the design approach. Systematizing abstraction between stages of the design approach with added deliverables such as requirements specifications (Balzert, 2010) or design principles and design features (Feine et al., 2020; B. Kuechler and Vaishnavi, 2008; W. Kuechler and Vaishnavi, 2012) can lead to better coordination and cooperation between stakeholders in a long-term project. This level of systematization supports the goal of GD research that allows the researchers in other contexts to reuse the knowledge from different projects to reconfigure or extend the developed artifacts in other research endeavors.

Broader traceability supporting iterative design

Rohde et al. (2017) alerted about the problem of self-referentiality in GD and pointed out the need to organize the design process to incorporate the growing understanding of the social practices because of the long-term research project.

“The design of IT artifacts is a reflexive endeavor in the sense that the artifacts’ appropriation and use change the organization’s social practices ... Design methodology must therefore cope with this inescapable fact and organize design and implementation processes in a reflexive and evolutionary way with iteratively revised and improved versions of the artifact each time leading to a new social practice” (Rohde et al. 2017, p. 166).

In this contribution, we second the point made by Rohde et al. and bring forward our practical experiences through two GD projects that a mechanism to forward and backtrack information flow between different iterations of analysis and design is an effective means to support the reflexive and evolutionary nature of GD research. The systematization between design approach stages with the induction of structured requirements elicitation methods from IS and BI as applied in the above-mentioned project exhibits invoke a sense of broader traceability in the GD research project. The traceability originates as the design features can be backtracked to analyze social practices and their evidence in empirical data and vice versa. As the project takes its course following iteration with the ongoing context study (examining users’ social practices), as well as the appropriation study (evaluating the prototype and analyzing the feedback to be incorporated into the design), the design study needs to include the knowledge about social practices coming from these multiple sources. The traceability originated as the blending of structured requirements elicitation practice between stages of the design approach can ease up the knowledge incorporation into the design. This traceability also offers a self-reflexive point of view on the research methods applied by the researchers by providing means to track decisions and rectify errors along the way.

Conclusion

This research presents the experiences of two GD research projects to expand the conceptual and practical understanding of GD research practices. It builds upon the notion of the evolutionary and iterative nature of design approaches in the context of GD and highlights the obstacles in dealing with growing knowledge of social phenomena during the research and design process. The incoming continuous knowledge needs incorporation in design and hence requires a systematized knowledge transition to accommodate changes in design iteratively. We suggest method appropriation from dissimilar fields of design like BI and software engineering to systematize knowledge between the analysis and design stages of GD projects. Through this contribution, we demonstrate that this systematization

of knowledge expands the horizon of GD research, affirming the flexibility and creative liberty promised by its praxeological framework. It further supports the abstraction in design approaches and addresses the methodological challenge of self-referentiality in long-term research projects.

References

- Balzert, H. (ed.) (2010): *Lehrbuch der Softwaretechnik: Basiskonzepte und Requirements Engineering*, Springer Spektrum, Heidelberg, <https://doi.org/10.1007/978-3-8274-2247-7>
- Bannon, L. J. (1995): 'From Human Factors to Human Actors: The Role of Psychology and Human-Computer Interaction Studies in System Design'. in R. M. Baecker, J. Grudin, W. A. S. Buxton and S. Greenberg (eds.), *Readings in Human-Computer Interaction*, 2nd ed., Kaufmann Morgan, San Francisco, 1995, pp. 205-214. <https://doi.org/10.1016/B978-0-08-051574-8.50024-8>
- Basil, V. R. and Turner, A. J. (1975): 'Iterative Enhancement: A Practical Technique for Software Development', *IEEE Transactions on Software Engineering*, vo. 1, no. 4, pp. 390-396. <https://doi.org/10.1109/TSE.1975.6312870>
- Baskerville, R., Baiyere, A., Gregor, S., Hevner, A. and Rossi, M. (2018): 'Design science research contributions: Finding a balance between artifact and theory', *Journal of the Association for Information Systems*, vol. 19, no. 5. <https://doi.org/10.17705/1jais.00495>
- Benyon, D. (2019): *Designing User Experience: A guide to HCI, UX and interaction design*, Pearson Education, London [u.a.].
- Bernhard, E., Recker, J. and Burton-Jones, A. (2013): 'Understanding the actualization of affordances: A study in the process modeling context', *Proceedings of the 34th International Conference on Information Systems (ICIS 2013)*, vol. 5, pp. 1-11.
- Braun, V. and Clarke, V. (2006): 'Using thematic analysis in psychology', *Qualitative Research in Psychology*, vol. 3, no. 2. <https://doi.org/10.1191/1478088706qp063oa>
- Brödner, P. (2009): 'The misery of digital organisations and the semiotic nature of IT, AI and Society', *Journal of Knowledge, Culture and Communication*, vol. 3, no. 3, pp. 331-351. <https://doi.org/10.1007/s00146-006-0066-1>
- Brown, T. (2008): 'Design thinking', *Harvard Business Review*, vol. 86, no. 6, pp. 84-92. Retrieved February 20, 2020 from: <https://hbr.org/2008/06/design-thinking>
- Chandra, L., Seidel, S. and Gregor, S. (2015): 'Prescriptive knowledge in IS research: Conceptualizing design principles in terms of materiality, action, and boundary conditions', *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2015, pp. 4039-404, <https://doi.org/10.1109/HICSS.2015.485>
- Diaper, D. and Stanton, N. A. (2004): *The Handbook of Task Analysis for Human-Computer Interaction*. Lawrence Erlbaum Associates Publishers, Mahwah, New Jersey.
- Dijkstra, E. W. (1968): 'The Structure of THE-Multiprogramming System', *Communications of the ACM*, vol. 11, no. 5, pp. 341-346. <https://doi.org/10.1145/357980.357999>
- Dourish, P. (2003): 'The appropriation of interactive technologies: Some lessons from placeless documents', *Computer Supported Cooperative Work: CSCW: An International Journal*, vol. 12, no. 4, pp. 465-490. <https://doi.org/10.1023/A:1026149119426>
- Emerson, R. M., Fretz, R. I. and Shaw, L. L. (2011): *Writing ethnographic fieldnotes*, 2nd ed., Univ. of Chicago Press, Chicago.
- Feine, J., Morana, S. and Maedche, A. (2020): 'Designing interactive chatbot development systems', *International Conference on Information Systems (ICIS 2020) - Making Digital Inclusive: Blending the Local and the Global*, India: AISel.
- Flick, U. (2018): *The SAGE Handbook of Qualitative Data Collection*, SAGE Publications Ltd., London. <https://doi.org/10.4135/9781526416070>
- Frost, N. (2016): *Practising Research. Why you're always part of the research process even when you think you're not*, Bloomsbury Publishing, London, U.K.

- Frost, N., Nolas, S., Brooks-Gordon B., Esin C., Holt A., Mehdizadeh L. and Shinebourne, P. (2010): 'Pluralism in qualitative research: The impact of different researchers and qualitative approaches on the analysis of qualitative data', *Qualitative Research*, vol. 10, no. 4, pp. 441-460.
- Gaver, W. (2012): 'What should we expect from research through design?', *Conference on Human Factors in Computing Systems – Proceedings (CHI 2012)*, pp. 937-946. <https://doi.org/10.1145/2207676.2208538>
- Goldkuhl, G. (2012): 'Pragmatism vs interpretivism in qualitative information systems research', *European Journal of Information Systems*, vol. 21, no. 2, pp. 135-146. <https://doi.org/10.1057/ejis.2011.54>
- Gregor, S. (2006): 'The nature of theory in Information Systems', *MIS Quarterly: Management Information Systems*, vol. 30, no. 3, pp. 611-642. <https://doi.org/10.2307/25148742>
- Harper, R., Rodden, T., Rogers, Y. and Sellen, A. (2008): *Being Human: Human-Computer Interaction in the year 2020*, Microsoft Research Ltda., Cambridge, England.
- Hatterscheid, E. and Schluter, F. (2018): 'Towards a decision support approach for selecting physical objects in collaborative supply chain processes for cyber physical system-transformation', *5th International Conference on Industrial Engineering and Applications, ICIEA 2018*, pp. 1-5. <https://doi.org/10.1109/IEA.2018.8387062>
- Hoare, C. A. R. (1971): 'Proof of a program: FIND', *Communications of the ACM*, vol. 14, no. 1, pp. 39-45. <https://doi.org/10.1145/362452.362489>
- Hoök, K. and Lowgren, J. (2012): 'Strong concepts: Intermediate-level knowledge in interaction Design research', *ACM Transactions on Computer-Human Interaction*, vol. 19, no. 3, pp. 1-18. <https://doi.org/10.1145/2362364.2362371>
- Holtzblatt, K. and Jones, S., (2017): 'Contextual Inquiry: A Participatory Technique for System Design'. in D. Schuler and A. Namioka. (eds.): *Participatory Design: Principles and Practices*, Erlbaum, Hillsdale, NJ [u.a.], DOI 10.1201/9780203744338-9.
- Kuckartz, U. and Rädiker, S. (2019): *Analyzing Qualitative Data with MAXQDA*. Springer International Publishing, New York [u. a.]. <https://doi.org/10.1007/978-3-030-15671-8>
- Kuechler, B. and Vaishnavi, V. (2008): 'On theory development in design science research: Anatomy of a research project', *European Journal of Information Systems*, vol. 17, no. 5, pp. 489-504. <https://doi.org/10.1057/ejis.2008.40>
- Kuechler, W. and Vaishnavi, V. (2012): 'A framework for theory development in design science research: Multiple perspectives', *Journal of the Association for Information Systems*, vol. 13, no. 6. <https://doi.org/10.17705/1jais.00300>
- Kuechler, W., Vaishnavi, V. K. and Petter, S. (2005): 'The Aggregate General Design Cycle as a Perspective on the Evolution of Computing Communities of Interest', *Computing Letters*, vol. 1, no. 3, pp. 123-128. <https://doi.org/10.1163/1574040054861221>
- Kuutti, K. and Bannon, L. J. (2014): 'The turn to practice in HCI: Towards a research agenda', *Proceedings of the Conference on Human Factors in Computing Systems (CHI 2014)*, pp. 3543-3552. <https://doi.org/10.1145/2556288.2557111>
- Lohrmann, M. and Reichert, M. (2016): 'Effective application of process improvement patterns to business processes', *Software and Systems Modeling*, vol. 15, no. 2, pp. 353-375. <https://doi.org/10.1007/s10270-014-0443-z>
- Luqi (1989): 'Software Evolution Through Rapid Prototyping', *IEEE Computer*, vol. 22, no. 5. pp. 13-25. <https://doi.org/10.1109/2.27953>
- MacCormack, A., Verganti, R. and Iansiti, M. (2001): 'Developing products on 'Internet time': The anatomy of a flexible development process', *Management Science*, vol. 47, no. 1, pp. 133-150. <https://doi.org/10.1287/mnsc.47.1.133.10663>
- Majchrzak, T. A., Sakurai, M. and Serrano, N. (2018): 'Conceptualizing and designing a resilience information portal', *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2018-January, pp. 45-54, <https://doi.org/10.24251/hicss.2018.008>
- Mao, J. Y., Vredenburg, K., Smith, P. W., & Carey, T. (2005): 'The state of user-centered design practice', *Communications of the ACM*, vol. 48, no. 3, pp. 105-109.
- Mauthner, N. and Doucet, A. (2003): 'Reflexive accounts and accounts of reflexivity in qualitative data analysis', *Sociology*, vol. 37. No. 3, pp. 413-431.

- McDonald, S. (2005): 'Studying actions in context: A qualitative shadowing method for organizational research', *Qualitative Research*, vol. 5, no. 4, pp. 455-473. <https://doi.org/10.1177/1468794105056923>
- Muller, M. J. and Kuhn, S. (1993): 'Participatory design', *Communications of the ACM*, vol. 36, no. 6, pp. 24-28. <https://doi.org/10.1145/153571.255960>
- Mydin, F. and Surat, S. (2021): 'Research capability: Early-career academics' perception of doctoral studies', *European Journal of Educational Research*, vol. 10, no. 4, pp. 1893-1905. <https://doi.org/10.12973/EU-JER.10.4.1893>
- Quinlan, E. (2008): 'Conspicuous invisibility: Shadowing as a data collection strategy', *Qualitative Inquiry*, vol. 14, no. 8, pp. 1480-1499. <https://doi.org/10.1177/1077800408318318>
- Rohde, M., Brödner, P., Stevens, G., Betz, M. and Wulf, V. (2017): 'Grounded Design - A praxeological IS research perspective', *Journal of Information Technology*, vol. 32, no. 2, pp. 163-179. <https://doi.org/10.1057/jit.2016.5>
- Royce, W. W. (1970): 'Managing the Development of Large Software Systems', *Proceedings of IEEE WESCON*, vol. 26, pp. 1-9.
- Rüller, S., Aal, K., Holdermann, S., Tolmie, P., Hartmann, A., Rohde, M., ... & Wulf, V. (2021): 'Technology is Everywhere, we have the Opportunity to Learn it in the Valley': The Appropriation of a Socio-Technical Enabling Infrastructure in the Moroccan High Atlas. Computer Supported Cooperative Work (CSCW), 1-40. <https://doi.org/10.1007/s10606-021-09401-8>
- Seidel, S., Chandra Kruse, L., Székely, N., Gau, M. and Stieger, D. (2018): 'Design principles for sensemaking support systems in environmental sustainability transformations', *European Journal of Information Systems*, vol. 27, no. 2, <https://doi.org/10.1057/s41303-017-0039-0>
- Self, B. (2021): 'Conducting Interviews During the COVID-19 Pandemic and Beyond', *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, vol. 22, no. 3. <https://doi.org/https://doi.org/10.17169/fqs-22.3.3741>
- Sharp, H., Rogers, Y. and Preece, J. (2019): *Interaction Design: Beyond Human-Computer Interaction*, John Wiley & Sons, Indianapolis.
- Silverman, D. (2011): *Doing qualitative research: a practical handbook*, 3rd edition, Sage Publications, Los Angeles [u.a.].
- Stevens, G., Pipek, V. and Wulf, V. (2010): 'Appropriation Infrastructure: Mediating appropriation and production work, special issue on end user development', *Journal of Organizational and End User Computing*, vol. 22, no. 2, pp. 58-81. <https://doi.org/10.4018/JOEUC.2010040104>
- Stolterman, E. and Wiberg, M. (2010): 'Concept-driven interaction design research', *Human-Computer Interaction*, vol. 25, no. 2, pp. 95-118. doi.org/10.1080/07370020903586696
- Strauss, A. L. and Corbin, J. M. (1998): *Basics of qualitative research: techniques and procedures for developing grounded theory*, Sage Publications, Los Angeles [u.a.].
- Wulf, V.; Pipek, V.; Randall, D.; Rohde, M.; Schmidt, K. and Stevens, G. (eds.) (2018): *Socio Informatics - A Practice-based Perspective on the Design and Use of IT Artefacts*, Oxford University Press, Oxford.
- Wulf, V., Müller, C., Pipek, V., Randall, D., Rohde, M. and Stevens, G. (2015): *Practice-Based Computing: Empirically Grounded Conceptualizations Derived from Design Case Studies*, Springer, London, U.K. https://doi.org/10.1007/978-1-4471-6720-4_7
- Wulf, V., Rohde, M., Pipek, V., and Stevens, G. (2011): 'Engaging with practices: design case studies as a research framework in CSCW', *Conference on Computer supported cooperative work (CSCW '11)*. Association for Computing Machinery, New York, NY, USA, 505-512. <https://doi.org/10.1145/1958824.1958902>

Gabriela Marcu (2022): Toward Trauma-Informed Design of Behavioral Interventions: A Case Study on Classroom Management. In: Proceedings of the 20th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Notes, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_n05

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_n05

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

Toward Trauma-Informed Design of Behavioral Interventions: A Case Study on Classroom Management

Gabriela Marcu
University of Michigan School of Information
gmarcu@umich.edu

Abstract. Trauma theory can complement behavior change theory in the design of digital interventions by providing different insights into the social and contextual factors that influence a person's behavior. Two thirds of the population in the U.S. and Europe have experienced at least one traumatic event in their lives, and trauma can have lasting effects on an individual's functioning and behavior. A trauma-informed approach to design recognizes the prevalence of trauma and considers its potential effects on an individual, to shape interactions that more effectively meet their needs. To illustrate how trauma-informed design can enhance a human-centered design process, I present a case study of a long-term project that resulted in the deployment of multiple digital technologies for classroom management. Through a retrospective needs assessment drawing on trauma theory and trauma-informed practices, I identify new possibilities for intervention by reframing children's behaviors, shifting the focus of intervention to their environment, and addressing structural inequity.

Introduction

Use of behavior change theory is common in the design of digital interventions. A significant body of CSCW literature has focused on behavior change by targeting specific behaviors, toward improving health and well-being (e.g., for a review of personal informatics literature see Epstein et al., 2020). Behavior change can involve designing around such strategies as reinforcement, rewards, consequences, goal setting, improving self-efficacy, and learning from observing others. However, focusing on a person's individual behavior does not work for all types of interventions. Employing digital technologies to increase data and reflection on behavior may even lead to harmful rumination (Eikey et al., 2021). The importance of designing for the collaborative, ecological, cultural, social, and emotional aspects surrounding a person's behavior has increasingly been highlighted (Chung et al., 2016; Kaziunas et al., 2017; Murnane et al., 2018; Marcu and Spiller, 2020; Figueiredo et al., 2018; Berry et al., 2021; Martin-Hammond and Purnell, 2022). Complementary theoretical approaches are needed that offer insight into the complex social and contextual factors that impact a person's behavior, which may call for designing different types of interventions.

In this paper, I argue that trauma theory and trauma-informed practices can add a useful perspective during the human-centered design process, to help us design effective behavioral interventions. I focus especially on the early stages of needs assessment, discovery, problem definition, and ideation that defines what role the digital intervention can and should play in addressing the problem as defined. Approximately two thirds of the population in the U.S. and Europe have experienced at least one traumatic event in their lives (Trautmann and Wittchen, 2018; Solomon and Davidson, 1997). A portion of those with trauma exposure will have such a severe reaction that they develop post-traumatic stress disorder (PTSD), but many will have a reaction that does not rise to the level of such a diagnosis, even as it has an adverse effect on their lives. Despite the existence and effectiveness of treatment options after trauma exposure, many individuals may not be conscious of how trauma has affected them, may not seek help, and when they do it is often not obvious that trauma is the underlying cause of their presenting challenges. Substance use, mental health conditions, chronic illness, housing instability, and incarceration have all been linked to trauma. Trauma-informed practice emerged from treatment and services for substance use, following the realization that focusing on substance use behavior would not lead to change without addressing trauma if it is at the root of that behavior. Trauma-informed behavior change theory is also being developed to show how "trauma impacts an individual's ability to undertake beneficial behavioural change" (Marks et al., 2022).

I discuss trauma-informed design within the school context for several reasons. Having originated within services for substance use and mental health, trauma-informed practices have expanded to schools and other settings in which clinical treatment is not the focus, but in which acknowledging the prevalence of

trauma enables the design of services that better meet the needs of individuals (Overstreet and Chafouleas, 2016; Elliott et al., 2005). Trauma-informed service delivery can be especially beneficial preventatively for health promotion, and “schools represent an opportune system for prevention and early intervention across domains related to child success” (Chafouleas et al., 2016). The school day provides a consistent routine for trauma-informed interventions to help students develop healthy social behavior (Perry and Daniels, 2016).

Schools are also a context in which digital interventions have been designed and explored using a range of approaches, from behavior change theory (Matic et al., 2014; Spiller et al., 2019) to social and emotional skill learning (Slovák and Fitzpatrick, 2015). Digital interventions can address behavior at the individual and classroom levels, and teachers may seek them out when they find themselves with inadequate training, resources, or support for effectively managing behavior in the classroom (Beran, 2005; Mitchell and Arnold, 2004). There have been calls against digital technologies that enable surveillance, control, and bias in the care of children through datafication of their behaviors that is decontextualized from their lived experiences (Manolev et al., 2019; Cho et al., 2020; Lu et al., 2021a,b)—but alternative approaches to design are unclear.

This paper makes the following contributions:

- Presenting trauma theory and trauma-informed practices as a complementary approach to behavior change theory in the design of digital interventions.
- Defining trauma-informed design of digital behavioral interventions.
- Comparing outcomes from a typical human-centered design process to other directions that could have been identified through trauma-informed design.
- Describing trauma-informed alternatives to designing digital interventions that increase datafication and control in schools.

Related Work

Trauma Theory

Traumas can be naturally occurring, such as an earthquake, wildfire, or epidemic; they can be accidents such as a train derailment, structural collapse, or car accident; or they can be intentionally caused such as violence, childhood abuse or neglect, sexual assault, or racial discrimination (SAMHSA, 2014b). Traumas can be experienced at an individual level (e.g., intimate partner violence), at a community level (e.g., school shooting), or at a mass level (e.g., war). Traumas can affect a culture and be passed down to generations who did not experience them directly—such as colonialism, genocide, slavery, or the forced relocation of Native Americans (SAMHSA, 2014b).

The following definition of psychological trauma from the U.S. Substance Abuse and Mental Health Services Administration (SAMHSA) is most commonly used across the country when bridging trauma theory and practice:

Individual trauma results from an event, series of events, or set of circumstances that is **experienced by an individual as** physically or emotionally harmful or life threatening and that has **lasting adverse effects** on the individuals' functioning and mental, physical, social, emotional, or spiritual well-being. (SAMHSA, 2014; emphasis added)

The two points that I have emphasized in this definition are most relevant for how we design behavioral interventions, and they explain why trauma tends to be overlooked as a cause of behavior. The exact nature of events or circumstances are not as important as how an individual experiences them. When feeling threatened or overwhelmed, different people will respond to the same event in different ways. However an individual responds to the event (e.g., what are commonly referred to as fight or flight responses) is their brain and body's way of surviving the real or perceived threat. Any response is therefore natural and healthy in the moment the trauma is experienced. Unfortunately, the same response may show up and influence a person's future interactions because "trauma results in a fundamental reorganization of the way mind and brain manage perceptions" (van der Kolk, 2015). A person can unknowingly and unexpectedly be influenced by a past trauma at any time, potentially making their perceptions and behaviors inappropriate or ineffective in the present moment.

Long-term adverse affects of trauma can also be prevented or reduced. Trauma can be viewed as a psychological wound, because the appropriate environment and circumstances can help an individual heal (van der Kolk, 2015). A trauma-informed approach therefore provides a theoretical framework and practical guidance for how individuals, programs, services, and organizations can contribute to cultures that are physically and emotionally safe and empowering for all (Esaki et al., 2013; Bloom and Farragher, 2013; SAMHSA, 2014a). To this end, trauma-informed practices aim to resist retraumatization:

The term [retraumatization] not only refers to the effect of being exposed to multiple events, but also implies the process of reexperiencing traumatic stress as a result of a current situation that mirrors or replicates in some way the prior traumatic experiences (e.g., specific smells or other sensory input; interactions with others; responses to one's surroundings or interpersonal context, such as feeling emotionally or physically trapped). (SAMHSA, 2014b)

Retraumatization often occurs in invisible and subtle ways such as microaggression or a lack of transparency, because a traumatic experience makes someone feel disempowerment and a loss of control (SAMHSA, 2014b). The purpose of trauma-informed practices is to create environments, services, and interactions that minimize the risk of retraumatization that could be caused by stress, unpredictability, misdiagnosis, conformity, coercion, seclusion, exclusion, discrimination, oppression, or erosion of cultural and personal identity (SAMHSA, 2014b; Bloom and Farragher, 2013; Comas-Díaz et al., 2019; Elliott et al., 2005).

Trauma-Informed Design as Building the Right Thing

Understanding the prevalence of trauma and its effects on behavior can lead to more effective intervention by enabling us to consider trauma as a potential factor through each decision in the design process. The scope of this paper is to address the earliest stages of design because our decisions therein set the course for the rest of an intervention's design. A trauma-informed approach to design must begin from the earliest stages of discovery, learning about current practices, assessing needs, and defining the problem to be solved. Determining the role of a digital intervention in addressing the problem requires ensuring that we are building the right thing, before we build the thing right.

In human-centered design, a trauma-informed approach can have a similar influence as when we consider accessibility. Designing with an understanding of how disabilities effect a person's lived experience informs designs that are ultimately more accessible and impactful for all. Similarly, the aim of trauma-informed practice is not to provide clinical treatment that directly addresses the trauma. Rather, trauma-informed practice can help us avoid isolating and decontextualizing a person's behavior to the point of losing sight of their humanity, or the broader context and lived experience wherein their true challenges lie. If trauma is the root cause of a person's challenges, and seeking treatment for PTSD or related symptoms could be most helpful for improving their health and well-being, they may need help in identifying and considering this course of action. Another aim, whether a person is or may eventually be pursuing such treatment, is to avoid inadvertently causing harm by encouraging behavior change that is fruitless if its cause is tied to trauma.

Applying Trauma-Informed Principles to Design

A trauma-informed approach must be tailored to each context, but is guided by a set of principles that are used to assess, implement, and evaluate an organization, service, or intervention's responsiveness to the adverse effects of trauma (SAMHSA, 2014a). There are six key principles of a trauma-informed approach: (1) safety, (2) trustworthiness and transparency, (3) peer support, (4) collaboration and mutuality, (5) empowerment, voice and choice, and (6) cultural, historical, and gender issues (SAMHSA, 2014a).

When applied to behavioral health interventions, these principles are adapted as follows (SAMHSA, 2014b):

- Recognize that trauma-related symptoms and behaviors originate from adapting to traumatic experiences
- View trauma in the context of individuals' environments
- Create a safe environment through consistency, transparency, and compassion
- Identify recovery from trauma as a primary goal
- Support control, choice, and autonomy
- Create collaborative relationships and participation opportunities

- Conduct universal routine trauma screening
- View trauma through a sociocultural lens
- Use a strengths-focused perspective to promote resilience

My aim is to discuss how these principles can be translated into the design of digital behavioral interventions. Trauma-informed design is a new concept as applied to digital technologies. The term began appearing in blog posts written by user experience (UX) professionals in the U.S., U.K., and Australia in 2021, which briefly discussed applying trauma-informed principles in human-centered design and interaction design projects¹²³. Trauma-informed design has perhaps gained more momentum in architecture, where, also in 2021, scholarship focused on the built environment began using the term trauma-informed design and describing its applicability to promoting the dignity, personal autonomy, and healing of individuals experiencing housing instability (Bollo and Donofrio, 2021). Trauma-informed design of school buildings and environments has similarly been identified as a critical part of trauma-informed schools (Kopec and Harte, 2020). This paper extends trauma-informed design to digital technologies and how they are conceived, generated, and deployed via the human-centered design process. In doing so, this paper advances the nascent scholarship on trauma in UX, for example (Chen et al., 2022) suggesting the application of trauma-informed principles to user research and user interface design.

Classroom Management

Classroom management refers to the strategies and behavior that a teacher uses to manage the individual and group behavior of their students (Emmer and Stough, 2001). Classroom management includes a broad range of evidence-based practices for: organizing and managing students as a group to help them stay on-task, fostering positive teacher-student and peer relationships, responding to the psychological and learning needs of individual students, and using counseling and behavioral methods to support students who demonstrate behavior problems (Emmer and Stough, 2001; Emmer and Sabornie, 2015).

A systematic review of digital technologies for classroom management classified them into three roles: (1) teacher training through video-based professional learning and interactive simulations, (2) tools for teachers and administrators such as classroom management apps used to give student points for or gamify their behavior, and (3) supporting student self-regulation by enabling them to log their own behavior and/or reflect on documentation of their behavior (Cho et al., 2020).

This recent review found that, of the three paradigms used for classroom management—behavioral, ecological, and social-emotional—digital technologies were most likely to support a behavioral approach (Cho et al., 2020). Moreover,

¹ <https://uxmag.com/articles/trauma-informed-design-understanding-trauma-and-healing>

² <https://uxdesign.cc/trauma-informed-design-research-69b9ba5f8b08>

³ <https://airbnb.design/designing-for-crisis/>

the use of digital technologies for collection and review of behavioral data by teachers and administrators has led to concerns that they can put more focus on data rather than the individual, promote a culture of control rather than care, and empower surveillance and bias rather than understanding (Manolev et al., 2019; Cho et al., 2020; Lu et al., 2021a,b). This paper extends this literature by proposing trauma-informed design of digital interventions as an approach that can help to avoid such unintended consequences, in part by shifting from a behavioral focus to the ecological and social-emotional.

Human-Centered Design Process

Over the course of a decade, I have studied behavioral intervention in school settings using the approaches of human-centered design and action research. I focused on special education services and supports for children in grades K-8 who display behavioral challenges. I began with fieldwork across settings observing daily interactions among children and staff, and broadly examining the role of digital technology in their activities. From informants such as educators, paraprofessional educators, therapists, social workers, psychologists, and clinical supervisors, I identified challenges in coordination of services. Together we designed and deployed several systems targeting these challenges, to learn more about digital technology's role in coordination, and identify strategies for how systems should be designed to meet sometimes competing stakeholder needs.

As I iteratively revisited problem definition and needs assessment, I continued down a path of how children's behaviors are documented so that appropriate services and accommodations can be provided to meet their needs. The intertwined systems of education and healthcare require staff to record behavioral data demonstrating each child's needs and progress, to account for additional levels of services when they are required (Marcu et al., 2021). The process I observed for recording, monitoring, and managing these data was a patchwork of paper data sheets, general software systems like spreadsheets, and specialized software systems such as those for reporting (Marcu et al., 2013). As a result, data collection emerged as a significant problem affecting the staff's work, with perceived consequences for the quality of care they felt they were able to provide children (Marcu et al., 2016). This problem was also corroborated by parents, who reported that they did not receive timely or detailed enough information to help them in understanding their child's behavior at school, the adequacy of services provided, or how their behavioral reinforcement at home might be aligned with the efforts of school staff (Marcu et al., 2019; Richards et al., 2021).

In the design and implementation stages of this work, I facilitated the successful adoption of technologies that became integrated into classroom settings because they met the needs of staff on the front lines of delivering behavioral services. I created a tablet-based data collection platform, Lilypad, which has been used in eight classrooms across three school districts. The longest deployment of three years has been with all of the staff in a three-classroom, self-contained

special education program focused on behavior disorders in a public elementary school. This program came to rely on Lilypad, requesting to use it for three consecutive school years until the COVID-19 pandemic upended their activities. They were persistent through the logistical and technical challenges of maintaining a research prototype for such a long period of time because they were so motivated to abandon paper-based data collection, which they referred to as being “back in the Stone Age” after adopting Lilypad. Our deployment of Lilypad also spread to additional schools by word of mouth when the staff from this program mentioned it to colleagues across their school district. Over time, I refined the Lilypad system by responding to user requests and adding various administrative features for customized data monitoring.

Additional design opportunities came from my observation of students’ tendency to seek updates on whether they were meeting behavioral expectations, along with the challenges that staff faced in providing them with consistent, individualized verbal feedback. Data management tools, including Lilypad, were largely focused on documenting behaviors for organizational use rather than using that data to provide meaningful feedback. Ideation and prototyping included individual desk-mounted displays and smartwatch applications, which could facilitate more frequent and individualized feedback for students, as well as opportunities to self-assess and self-monitor their own behavior at school. My most extensive exploration aligned with existing technologies, such as ClassDojo, which allow for class-wide behavior monitoring. Partnering with a teacher who had experience using ClassDojo but wanted a more custom form of data monitoring and feedback, I co-designed a wall-mounted display that enabled students to see their Lilypad data in real time throughout the school day. In a deployment lasting a full school year, I focused on classroom dynamics such as how the teacher incorporated the display into their activities, and how students reflected on the data collectively (Spiller et al., 2019).

The Limitations of Human-Centered Design

The human-centered design process enabled me to define and then meet the needs of users who were central in the delivery of services to children, toward how they wanted to improve the quality of care and behavioral outcomes for those children. I not only involved the users themselves in design and evaluation of these systems, but also domain experts who helped me design in support of evidence-based behavioral interventions. The resulting systems saw such metrics of success as adoption, acceptance, and sustained use over months or years.

Yet, I began to wonder whether behavioral data collection—encompassing the processes, practices, and tools I had homed in on through this process—may ultimately be incompatible with my goal of supporting coordination and my intended impact of improving outcomes for children. I found that behavioral data are subjective and difficult to represent within organizational records that might be used for coordination, such as a behavioral report card in schools (Marcu et al.,

2021). Moreover, paper-based and electronic records are created primarily for internal use, and as such do not meaningfully support distributed coordination (Marcu et al., 2016; Richards et al., 2021). Due to such issues, I identified record keeping as a factor contributing to breakdowns that cause stakeholders such as parents and teachers to be at odds with one another rather than working together (Marcu et al., 2019). My design efforts had met the needs of those primarily responsible for data collection, by allowing them to more easily capture, manage, and share data, while freeing up more of their attention for working with students and other stakeholders. But by maintaining the status quo and supporting existing practices centered around data, I did not explore design for more transformative change.

For example, an ongoing challenge throughout this work has been the difficulty of translating evidence-based strategies into classroom practice. According to the literature, punishing undesired behavior is not as effective as proactively setting clear expectations for behavior and then reinforcing behaviors that meet those expectations. But in real classroom settings, punishment is common, from verbal corrections to consequences such as loss of privileges. During iterative design and deployment of Lilypad, I explored this issue with users. We discussed how the Lilypad interface and functionality might reduce use of punishment by making it easier for staff to communicate expectations, positive reinforcement, and individualized feedback to students. However, these aspects of Lilypad were difficult to sustain and regularly abandoned.

Shifting to explore a broader role for technology beyond record keeping, I then compared practices among stakeholders to model how shared understanding of a child's needs and progress is developed across organizational boundaries, and described how we can design coordination tools that facilitate this process (Richards et al., 2021). A key finding was the importance of relationship building practices – often a distinct mechanism separate from data collection, and not always improved by the presence of digital technology. Having learned that Lilypad, strictly as a data collection or record keeping tool, could not improve coordination or prevent breakdowns, I still did not have a clear path for design alternatives (Richards et al., 2021).

Finally, I found a trauma-informed approach to be pivotal in helping me identify alternative ways that technologies like Lilypad could support students and staff. Interestingly, during a field visit early on, I happened upon an artifact that revealed the school's use of a trauma-informed approach, but at the time I did not understand its significance, and trauma-informed practice never came up over the course of many months working together. This could be because my informants did not think to bring up trauma-informed practices, perhaps thinking they were not relevant to the use of digital technologies or data collection. Or it could be because I did not know enough about trauma-informed practice at the time, or I did not ask the right questions of my informants. Whatever the case, I believe this missed opportunity speaks to the need for human-centered designers to be knowledgeable about trauma

theory and practice, as perhaps the only way to bring a trauma-informed approach to their own work.

This paper, and my definition of trauma-informed design, is the result of my reviewing literature and receiving training on trauma theory and trauma-informed practice. The following is a needs assessment that I conducted on this project mostly retrospectively, in an iterative, reflexive process over several years as I looked back on my design process the more I gained knowledge on trauma. This included comparing prior analyses across studies (e.g., interviewing vs. observational methods, Lilypad vs. ClassDojo use, teacher vs. parent perspectives) and applying trauma-informed principles to various junctions throughout this long-term project to assess how they would have influenced my interpretation or design decisions.

Trauma-Informed Design of Behavioral Interventions

A trauma-informed approach to design helped me consider how the effects of trauma could be influencing the behaviors of children as well as staff. Retrospectively performing a needs assessment through the lens of trauma changed my understanding of the issues, and also indicated different opportunities for design.

Understanding the Root Cause of Behavior

Children's behavior cannot be effectively addressed without an understanding of its root cause, which requires adequate training, tools, and time to assess. A trauma-informed approach covers one type of root cause, reframing problem behavior as potentially being a part of how students have been affected by adverse events. A person's brain, body, and behavior can undergo fundamental changes in response to a traumatic experience. These changes are a natural response to something very stressful, and are functional in allowing the person to survive and cope with the experience. After the traumatic experience, however, these changes may reappear in situations in which they are not functional. In the present, a person's behavior may appear on the surface to be inappropriate to the situation, but this is not something they can control even if this is pointed out to them or they face consequences. If the behavior stems from a past trauma, the best way to help the person adjust their behavior is by helping them process and heal from the trauma. Trauma can also be overlooked as a factor in assessing behavioral needs and misdiagnosing disabilities, for example "extra caution needs to be taken when children with the history of trauma are evaluated for ADHD, as the overlap with the symptoms of PTSD and other trauma-related psychopathology can easily lead to misdiagnosis" (Szymanski et al., 2011).

What does this mean for the role of a system like Lilypad in addressing student behavior? When we interviewed teachers with experience using classroom management technologies (ClassDojo is the most popular commercially available

app), they pointed out that their standard approach of quantifying a behavior and using gamification to increase or decrease a student's points in response to specific behaviors may not be address the root cause:

“If you have challenging students in the classroom, you know they're not on medication, you suspect that they might be ADHD or ADD. [...] But you're penalizing them for something that is beyond their control. So, that's another reason why I really don't like ClassDojo, [...] because teachers are going to end up punishing a child that can't help it. It's beyond their control. So why bother? I mean, this is not going to help.” (Lu et al., 2021b).

Tools like Lilypad and ClassDojo support school staff in their typical practices of documenting problematic student behavior by capturing quantitative data such as frequency and duration of behaviors, and organizing the data under broad categories such as not being safe, kind, or on task (Marcu et al., 2021). Enhancing these practices through streamlined data management and visualization increases the potential for staff to focus on behaviors noticeable on the surface, without probing for their root cause.

Identifying a past trauma that may be affecting a child's behavior can fundamentally shift the ways that staff respond to the behavior, and change the trajectory of how they plan behavioral interventions for the child. Trauma training and screening are the recommended ways to integrate an awareness of trauma into behavioral health services (von der Embse et al., 2019). Lilypad could support training all staff to understand the effects of trauma, and remind them to consider these effects each time they encounter various student behaviors. Lilypad could facilitate trauma screening via validated questionnaires (e.g., Lang and Connell; Bernstein et al., 2018), some of which are already developed for computerized or online administration (e.g., Frewen et al., 2013). Data collection tasks could be presented alongside relevant resources, so that users have materials on trauma and a trauma screening questionnaire on hand—serving as reminders to compare behaviors they are observing in students to what they have learned about trauma. A challenge of implementing trauma-informed practices is integrating them into all aspects of work across an organization or program. Thus, Lilypad could provide a structure for training staff, refreshing their training periodically, and ensuring screening occurs systematically and equitably for students without relying as much on staff to remember, prepare, or facilitate screeners.

Intervening on the Environment Rather Than the Individual

Another aspect of considering the effects of trauma on an individual is examining the interactions between them and their environment. Staff communicating their behavioral expectations and reinforcing them in certain ways may work for some students, while inadvertently causing retraumatization for other students who may be reminded of past traumatic experiences. Retraumatization occurs unexpectedly when stimuli in the person's environment causes them to feel as if they are re-experiencing a past traumatic event in their brain and body, rendering them

unable to focus on and respond to the present moment. This phenomenon can cause more problematic behavior and worsen interactions, for example if a student is called out for not following directions and the student responds by inappropriately talking back to the teacher. Developmental trauma, which stems from cumulative experiences of abuse or neglect in early childhood, can alter interactions with others (van der Kolk and Courtois, 2005). Racial trauma is caused by repeated exposure to race-based discrimination and devaluation, and is experienced at both the individual and collective levels, which can lead to internalized oppression and the wound of rage (Comas-Díaz et al., 2019; Hardy, 2013).

Understanding that behavior may be explained by different kinds of trauma can drastically change how environments respond to that behavior, and how staff determine appropriate interventions: “Youth of color are often prescribed anger management interventions, while rage from the hidden wound of racial oppression remains unaddressed” (Hardy, 2013). It is also important to confront the fact that the very institutions attempting to serve children—such as education, special education, healthcare, and behavioral research—also have a history of traumatizing children through systemic oppression, exclusion, and violence (Henderson et al., 2019; Marks et al., 2022; Powell, 2020; Yoon, 2019; Gagnon et al., 2017). This history influences interactions with a child and their community, requiring trust to be rebuilt over time, and opportunities for healing rather than perpetuating harms such as blaming indigenous parenting practices when assessing children’s behavior (Maxwell, 2014), or continuing use of practices like seclusion and restraint, which have been shown to be dangerous (Gagnon et al., 2017).

To shift the focus of interventions from the individual to their environment, the human-centered design process for a system like Lilypad could support classroom management by aiming to change the behavior of teachers and staff, rather than the behavior of students. Trauma-informed practices include providing classroom management training to school staff so that they cultivate an effective environment for their students (von der Embse et al., 2019). Use of Lilypad made some practices more visible so that perhaps staff could more easily reflect on them, but was not an effective tool for meaningfully changing staff’s behavior with children (Marcu et al., 2021). For example, Lilypad enabled staff to capture instances of a student using inappropriate language or exhibiting violent behavior toward peer, staff, or property. But key contextual details were not consistently documented, and there was no way to compare the student’s behavioral data to specific preventative or reactive strategies used by the staff. Some of the earliest—and most successful—Lilypad adopters were recommended by community gatekeepers because they displayed generally effective classroom management along with rigorous data collection practices, which they were motivated to transition from paper-based to computerized methods. In contrast, staff who would later approach us about using Lilypad were characterized by greater challenges with classroom management, and they hoped Lilypad could help them create and implement a classroom management plan. Meeting this need would have required us to shift the

role and design of Lilypad toward providing staff with structure and guidance for developing their classroom management plan.

Attending to the impacts of trauma and following a trauma-informed approach therefore leads to a markedly different type of intervention. Rather than collecting data on individual student behaviors for the purpose of analyzing patterns in their behavior with limited context, designing an intervention may instead focus entirely on their context by supporting the training of staff on classroom management, ensuring they have an effective classroom management plan, and facilitating how they enact that plan.

Targeting Structural Inequity

Finally, interventions concerned with trauma must address structural and historical causes of inequity. Students of color and students with disabilities, for example, have not historically had equal access to education, and in myriad ways continue to experience exclusion and discrimination. If behavior management systems like Lilypad focus on quantifying behavior in a decontextualized way, they can replicate existing bias (Lu et al., 2021b). Data collection can also perpetuate oppression if its purpose is functionally as surveillance and control—on students as well as teachers themselves, who may bear pressure from school administrators or parents (Lu et al., 2021a). I observed indications of such unintended consequences with the use of Lilypad, and confirmed them in a broader interview study with teachers using the similar commercially available app, ClassDojo. These apps help to make behavioral expectations more visible, and provide feedback to students through gamification—students can see each time they earn or lose points based on how well they are following behavioral rules throughout the school day. As a result, students may feel their behavior is being more closely observed, documented, and rewarded or punished in real time. Similarly, teachers know that the data they capture on their students can be reviewed by others and provide only a limited perspective on their classroom management skills.

At the same time, data are important to a trauma-informed approach because they can help to identify students “who have been exposed to adverse events and may be at risk of or are already exhibiting effects of those experiences” (Chafouleas et al., 2016). As I have discussed, framing behavior using trauma as a lens can shape how the behavior is addressed. Data that help to contextualize behavior and guide a trauma-informed approach to intervening may therefore serve to provide staff an alternative perspective, counteracting implicit bias that is more likely to play a role when they draw their own interpretations of behavioral data. For example, a system like Lilypad could structure collection and representation of data in such a way that staff approach the data with a trauma-informed lens, to holistically consider what factors may be affecting attendance patterns, behaviors, and grades (Chafouleas et al., 2016). Behavioral data might also be represented within the context of a student’s interactions with others, so that the staff are accountable for their role in creating a supportive and inclusive school climate.

However, I have observed that use of Lilypad focused on accounting for one's work can interfere with effectively delivering services to children (Marcu et al., 2021). Therefore, accounting of one's work should be linked to training and support for teachers, rather than placing teachers under surveillance or scrutiny (Lu et al., 2021a). A trauma-informed approach to delivering services recognizes that the organizational culture and processes must be democratic and ensure that staff feel empowered in doing their work.

Conclusion

I have defined trauma-informed design of behavioral interventions by drawing on trauma theory and practice. Psychological trauma can be understood as a wound caused by any event experienced by an individual as physically or emotionally harmful, with lasting effects on how the person functions in the world. With the appropriate environment and support an individual can heal from trauma, otherwise lasting adverse affects can change their behavior, mental health, or physical health. The aim of trauma-informed practices is not to treat trauma, but they do enable the healing process to take place by preventing further harm. Due to the high prevalence of trauma, it is important to have an understanding of its potential impact on individuals when designing services and interventions for behavioral health. Using a case study of the Lilypad platform, which supports behavioral data collection for classroom management, I showed how a trauma-informed approach reframes a child's behavior through the lens of trauma, and broadens the focus of intervention beyond a specific behavior to the child's social, emotional, organizational, and cultural context. With the increasing use of digital technologies for behavior management and discipline in schools, trauma-informed design can help us find alternatives to the use of these technologies for datafication, punishment, surveillance, and control.

Acknowledgments

I am grateful to the collaborators, students, mentors, and participants who have contributed to each stage of this work and shaped my thinking over the years. This article is based upon work supported by the National Science Foundation under Grant No. IIS 1816319.

References

- Bassuk, E. L., R. E. Latta, R. Sember, S. Raja, and M. Richard (2017): 'Universal Design for Underserved Populations: person-centered, recovery-oriented and trauma informed'. *Journal of Health Care for the Poor and Underserved*, vol. 28, no. 3, pp. 896–914.
- Beran, T. (2005): 'A new perspective on managing school bullying: Pre-service teachers' attitudes'. *Journal of Social Sciences*, vol. 8, no. 3, pp. 43–49.

- Bernstein, D. P., J. A. Stein, M. D. Newcomb, E. Walker, D. Pogge, T. Ahluvalia, J. Stokes, L. Handelsman, M. Medrano, D. Desmond, and W. Zule (2003): 'Development and validation of a brief screening version of the Childhood Trauma Questionnaire'. *Child Abuse & Neglect*, vol. 27, no. 2, pp. 169–190.
- Berry, A. B., C. Y. Lim, C. A. Liang, A. L. Hartzler, T. Hirsch, D. M. Ferguson, Z. A. Bernet, and J. D. Ralston (2021): 'Supporting Collaborative Reflection on Personal Values and Health'. *Proc. ACM Hum.-Comput. Interact.*, vol. 5, no. CSCW2.
- Bloom, S. L. and B. Farragher (2013): *Restoring sanctuary: A new operating system for trauma-informed systems of care*. Oxford University Press.
- Bollo, C. and A. Donofrio (2021): 'Trauma-informed design for permanent supportive housing: four case studies from Seattle and Denver'. *Housing and Society*, pp. 1–22.
- Bowers, J. M. (1991): 'The Janus Faces of Design: Some Critical Questions for CSCW'. In: J. M. Bowers and S. D. Benford (eds.): *Studies in Computer Supported Cooperative Work: Theory, Practice and Design*. Amsterdam, etc., pp. 333–350, North-Holland.
- Chafouleas, S. M., A. H. Johnson, S. Overstreet, and N. M. Santos (2016): 'Toward a blueprint for trauma-informed service delivery in schools'. *School Mental Health*, vol. 8, no. 1, pp. 144–162.
- Chen, J. X., A. McDonald, Y. Zou, E. Tseng, K. A. Roundy, A. Tamersoy, F. Schaub, T. Ristenpart, and N. Dell (2022): 'Trauma-Informed Computing: Towards Safer Technology Experiences for All'. In: *CHI Conference on Human Factors in Computing Systems*. pp. 1–20.
- Cho, V., K. C. Mansfield, and J. Claughton (2020): 'The past and future technology in classroom management and school discipline: A systematic review'. *Teaching and Teacher Education*, vol. 90, pp. 103037.
- Chudzik, M., C. Corr, and K. Wolowiec-Fisher (2022): 'Trauma: Early Childhood Special Education Teachers' Attitudes and Experiences'. *Early Childhood Education Journal*, pp. 1–12.
- Chung, C.-F., K. Dew, A. Cole, J. Zia, J. Fogarty, J. A. Kientz, and S. A. Munson (2016): 'Boundary negotiating artifacts in personal informatics: patient-provider collaboration with patient-generated data'. In: *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*. pp. 770–786.
- Comas-Díaz, L., G. N. Hall, and H. A. Neville (2019): 'Racial trauma: Theory, research, and healing: Introduction to the special issue.'. *American Psychologist*, vol. 74, no. 1, pp. 1.
- Eikey, E. V., C. M. Caldeira, M. C. Figueiredo, Y. Chen, J. L. Borelli, M. Mazmanian, and K. Zheng (2021): 'Beyond self-reflection: introducing the concept of rumination in personal informatics'. *Personal and Ubiquitous Computing*, vol. 25, no. 3, pp. 601–616.
- Elliott, D. E., P. Bjelajac, R. D. Falloot, L. S. Markoff, and B. G. Reed (2005): 'Trauma-informed or trauma-denied: Principles and implementation of trauma-informed services for women'. *Journal of Community Psychology*, vol. 33, no. 4, pp. 461–477.
- Emmer, E. T. and E. J. Sabornie (2015): *Handbook of Classroom Management*. Routledge New York.
- Emmer, E. T. and L. M. Stough (2001): 'Classroom management: A critical part of educational psychology, with implications for teacher education'. *Educational psychologist*, vol. 36, no. 2, pp. 103–112.

- Epstein, D. A., C. Caldeira, M. C. Figueiredo, X. Lu, L. M. Silva, L. Williams, J. H. Lee, Q. Li, S. Ahuja, Q. Chen, P. Dowlatyari, C. Hilby, S. Sultana, E. A. Eikey, and Y. Chen (2020): 'Mapping and taking stock of the personal informatics literature'. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, no. 4, pp. 1–38.
- Esaki, N., J. Benamati, S. Yanosy, J. S. Middleton, L. M. Hopson, V. L. Hummer, and S. L. Bloom (2013): 'The sanctuary model: Theoretical framework'. *Families in Society*, vol. 94, no. 2, pp. 87–95.
- Figueiredo, M. C., C. Caldeira, E. V. Eikey, M. Mazmanian, and Y. Chen (2018): 'Engaging with health data: The interplay between self-tracking activities and emotions in fertility struggles'. *Proceedings of the ACM on Human-Computer Interaction*, vol. 2, no. CSCW, pp. 1–20.
- Frewen, P. A., B. Evans, J. Goodman, A. Halliday, J. Boylan, G. Moran, J. Reiss, A. Schore, and R. A. Lanius (2013): 'Development of a childhood attachment and relational trauma screen (CARTS): A relational-socioecological framework for surveying attachment security and childhood trauma history'. *European Journal of Psychotraumatology*, vol. 4, no. 1, pp. 20232.
- Gagnon, D. J., M. J. Mattingly, and V. J. Connelly (2017): 'The restraint and seclusion of students with a disability: Examining trends in U.S. school districts and their policy implications'. *Journal of Disability Policy Studies*, vol. 28, no. 2, pp. 66–76.
- Gerson, E. M. and S. L. Star (1986): 'Analyzing due process in the workplace'. *ACM Transactions on Office Information Systems*, vol. 4, no. 3, pp. 257–270.
- Greene, N. (2020): 'Misdiagnosing Generational Trauma and Grief: I Am Not Angry; I Am Triggered and Grief Stricken'. In: *Narratives of Hope and Grief in Higher Education*. Springer, pp. 139–152.
- Hardy, K. V. (2013): 'Healing the hidden wounds of racial trauma'. *Reclaiming Children and Youth*, vol. 22, no. 1, pp. 24.
- Henderson, D. X., L. Walker, R. R. Barnes, A. Lunsford, C. Edwards, and C. Clark (2019): 'A framework for race-related trauma in the public education system and implications on health for black youth'. *Journal of School Health*, vol. 89, no. 11, pp. 926–933.
- Johansen, R. (1988): *Groupware. Computer Support for Business Teams*. New York and London: The Free Press.
- Kaziunas, E., M. S. Ackerman, S. Lindtner, and J. M. Lee (2017): 'Caring through data: Attending to the social and emotional experiences of health datafication'. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. pp. 2260–2272.
- Kaziunas, E., M. S. Klinkman, and M. S. Ackerman (2019): 'Precarious interventions: Designing for ecologies of care'. *Proceedings of the ACM on Human-Computer Interaction*, vol. 3, no. CSCW, pp. 1–27.
- Kopec, D. and J. D. Harte (2020): 'Design as the missing variable in trauma-informed schools'. In: E. Rossen and R. Hull (eds.): *Supporting and educating traumatized students: A guide for school-based professionals*. Oxford University Press, pp. 343—357.
- Lang, J. M. and C. M. Connell (2018): 'The child trauma screen: A follow-up validation'. *Journal of traumatic stress*, vol. 31, no. 4, pp. 540–548.

- Lu, A. J., T. R. Dillahunt, G. Marcu, and M. S. Ackerman (2021a): ‘Data Work in Education: Enacting and Negotiating Care and Control in Teachers’ Use of Data-Driven Classroom Surveillance Technology’. *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW2, pp. 1–26.
- Lu, A. J., G. Marcu, M. S. Ackerman, and T. R. Dillahunt (2021b): ‘Coding Bias in the Use of Behavior Management Technologies: Uncovering Socio-technical Consequences of Data-driven Surveillance in Classrooms’. In: *Designing Interactive Systems Conference 2021*. pp. 508–522.
- Manolev, J., A. Sullivan, and R. Slee (2019): ‘The datafication of discipline: ClassDojo, surveillance and a performative classroom culture’. *Learning, Media and Technology*, vol. 44, no. 1, pp. 36–51.
- Marcu, G., A. K. Dey, and S. Kiesler (2014): ‘Designing for collaborative reflection’. In: *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*. pp. 9–16.
- Marcu, G., A. K. Dey, and S. Kiesler (2021): ‘Tensions in Representing Behavioral Data in an Electronic Health Record’. *Computer Supported Cooperative Work (CSCW)*, vol. 30, no. 3, pp. 393–424.
- Marcu, G., A. K. Dey, S. Kiesler, and M. Reddy (2016): ‘Time to reflect: Supporting health services over time by focusing on collaborative reflection’. In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. pp. 954–964.
- Marcu, G., A. Spiller, J. Arevalo Garay, J. E. Connell, and L. R. Pina (2019): ‘Breakdowns in home-school collaboration for behavioral intervention’. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. pp. 1–12.
- Marcu, G. and A. N. Spiller (2020): ‘Collaborative Aspects of Collecting and Reflecting on Behavioral Data’. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. pp. 1–13.
- Marcu, G., K. Tassini, Q. Carlson, J. Goodwyn, G. Rivkin, K. J. Schaefer, A. K. Dey, and S. Kiesler (2013): ‘Why do they still use paper? Understanding data collection and use in Autism education’. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. pp. 3177–3186.
- Marks, C., J. L. Pearson, M. L. Zúñiga, N. Martin, D. Werb, and L. R. Smith (2022): ‘Articulating the Trauma-Informed Theory of Individual Health Behavior’. *Stress and Health*, vol. 38, no. 1, pp. 154–162.
- Martin-Hammond, A. and T. S. Purnell (2022): ‘Bridging Community, History, and Culture in Personal Informatics Tools: Insights from an Existing Community-Based Heart Health Intervention for Black Americans’. *Proc. ACM Hum.-Comput. Interact.*, vol. 6, no. GROUP.
- Matic, A., G. R. Hayes, M. Tentori, M. Abdullah, and S. Schuck (2014): ‘Collective use of a situated display to encourage positive behaviors in children with behavioral challenges’. In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. pp. 885–895.
- Maxwell, K. (2014): ‘Historicizing historical trauma theory: Troubling the trans-generational transmission paradigm’. *Transcultural Psychiatry*, vol. 51, no. 3, pp. 407–435.
- Mitchell, A. and M. Arnold (2004): ‘Behavior Management Skills as Predictors of Retention Among South Texas Special Educators’. *Journal of Instructional Psychology*, vol. 31, no. 3.

- Murnane, E. L., T. G. Walker, B. Tench, S. Volda, and J. Snyder (2018): ‘Personal Informatics in Interpersonal Contexts: Towards the Design of Technology That Supports the Social Ecologies of Long-Term Mental Health Management’. *Proc. ACM Hum.-Comput. Interact.*, vol. 2, no. CSCW.
- Overstreet, S. and S. M. Chafouleas (2016): ‘Trauma-informed schools: Introduction to the special issue’. *School Mental Health*, vol. 8, no. 1, pp. 1–6.
- Perry, D. L. and M. L. Daniels (2016): ‘Implementing trauma—informed practices in the school setting: A pilot study’. *School Mental Health*, vol. 8, no. 1, pp. 177–188.
- Powell, T. M. (2020): ‘The Scars of Suspension: Testimonies as Narratives of School-Induced Collective Trauma’. Ph.D. thesis, University of California, Los Angeles.
- Richards, O. K., A. Choi, and G. Marcu (2021): ‘Shared Understanding in Care Coordination for Children’s Behavioral Health’. *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW1, pp. 1–25.
- SAMHSA (2014a): *SAMHSA’s Concept of Trauma and Guidance for a Trauma-Informed Approach*. Rockville, MD: Substance Abuse and Mental Health Services Administration. HHS Publication No. (SMA) 14-4884.
- SAMHSA (2014b): *Trauma-Informed Care in Behavioral Health Services*. Rockville, MD: Substance Abuse and Mental Health Services Administration. Treatment Improvement Protocol (TIP) Series 57. HHS Publication No. (SMA) 13-4801.
- Slovák, P. and G. Fitzpatrick (2015): ‘Teaching and developing social and emotional skills with technology’. *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 22, no. 4, pp. 1–34.
- Solomon, S. D. and J. R. Davidson (1997): ‘Trauma: prevalence, impairment, service use, and cost’. *Journal of Clinical Psychiatry*, vol. 58, no. 9, pp. 5–11.
- Spiller, A. N., K. Caro, J. A. Garay, and G. Marcu (2019): ‘Supporting behavior management with a classroom display providing immediate feedback to students’. In: *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*. pp. 159–168.
- Szymanski, K., L. Sapanski, and F. Conway (2011): ‘Trauma and ADHD—association or diagnostic confusion? A clinical perspective’. *Journal of Infant, Child, and Adolescent Psychotherapy*, vol. 10, no. 1, pp. 51–59.
- Trautmann, S. and H.-U. Wittchen (2018): ‘Trauma and PTSD in Europe’. *Post-Traumatic Stress Disorder*, pp. 133–145.
- van der Kolk, B. A. (2015): *The body keeps the score: Brain, mind, and body in the healing of trauma*. Penguin Books.
- van der Kolk, B. A. and C. A. Courtois (2005): ‘Editorial comments: Complex developmental trauma’. *Journal of Traumatic Stress*, vol. 18, no. 5, pp. 385–388.
- Veinot, T. C., H. Mitchell, and J. S. Ancker (2018): ‘Good intentions are not enough: how informatics interventions can worsen inequality’. *Journal of the American Medical Informatics Association*, vol. 25, no. 8, pp. 1080–1088.
- von der Embse, N., L. Rutherford, A. Mankin, and A. Jenkins (2019): ‘Demonstration of a trauma-informed assessment to intervention model in a large urban school district’. *School Mental Health*, vol. 11, no. 2, pp. 276–289.

Yoon, I. H. (2019): 'Haunted trauma narratives of inclusion, race, and disability in a school community'. *Educational Studies*, vol. 55, no. 4, pp. 420–435.

Zimmerman, J., J. Forlizzi, and S. Evenson (2007): 'Research through design as a method for interaction design research in HCI'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 493–502.