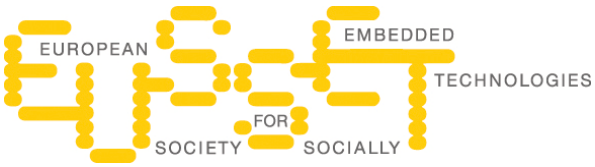
ISSN 2510-2591



Reports of the European Society for Socially Embedded Technologies

volume 3 issue 1
2019

Proceedings of 17th European Conference on Computer-Supported Cooperative Work - Exploratory Papers

Guest Editors

Pernille Bjørn (University of Copenhagen & University of Washington, USA)
Norman Makoto Su (Indiana University Bloomington, USA)

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/

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

Skarlatidou, A., Trimm, C., Vitos, M., Haklay, M. (2019): Designing Data Collection Interfaces for Low-literate Users. In: Proceedings of the 17th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing an the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep02

Designing Collaborative Data Collection Interfaces for Low-literate Users

Artemis Skarlatidou[1], Caroline Trimm[2], Michalis Vitos[1], Muki Haklay[1]

[1]Department of Geography, University College London; [2]UCL interaction Centre, University College London

a.skarlatidou@ucl.ac.uk, caroline.trimm.15@ucl.ac.uk, michalis.vitos@ucl.ac.uk, m.haklay@ucl.ac.uk

Abstract. Data collection applications on smartphone devices support indigenous communities in developing countries to record and preserve traditional ecological knowledge, collaboratively collect data around issues that are important to them and use these tools to subsequently identify locally-acceptable solutions with global impacts. Development of these interfaces needs to consider users' familiarity with technology as well as their education and literacy levels. This study builds on existing HCI4D research, which is also of interest to the CSCW community, in order to develop and evaluate, for their usability and user preferences, four user interfaces with low-literate people in the UK. Our findings suggest that linear navigation structures and a tangible interface are almost equally usable and preferred when they require minimum interaction with the device. Our preliminary analysis provides a deeper insight into the design issues to inform development of smartphone-based interfaces using various interaction types and we report on our methodological challenges from carrying out HCI research with low-literate people in the UK. The findings of this paper are used to inform the experimental design of additional work that we carry out with low-literate users in Namibia.

Introduction

Western beliefs that techno-scientific innovation, complex legislation, international agreements and top-down approaches can provide the solution and let us live a sustainable future have started slowly to fall apart. This is due to the widely documented disconnect that these strategies have from their actual recipients. Jerome Lewis who works with pygmy hunter-gatherers, explains that "people are integral to how their environments are shaped and the diversity that these environments support" (SynchronicityEarth.org, 2018). Excluding local communities from the broader sustainability debate and agenda not only disconnects us from primary sustainability goals but this further leads into strategies that are doomed to create unsustainable solutions.

For thousands of years people had to rely on their local environments to satisfy basic needs and through time communities have developed significant knowledge to help them deal with local issues. Amongst other types of knowledge, traditional ecological knowledge (TEK), is recognised within indigenous communities for millennia and it started to receive some attention from western knowledge structures and paradigms for its potential to support local and global sustainability. In line with inclusion and the 'leaving no one behind' principles of the UN's 2030 agenda for sustainable development, this requires zooming into local environments and their people to understand how they interact with them. With that aim in mind Extreme Citizen Science (ExCiteS) is a philosophy of

> "situated, bottom up practices which take into account local needs, practices and cultures and which work with broad networks of people in order to design and build new devices as well as knowledge creation processes which can truly transform the world".

Central to this philosophy are collaborative data collection tools, which support individuals and communities in the collection of knowledge they choose to preserve or in the collection of evidence which helps them demonstrate their local issues, an essential requirement in order to subsequently take further action which may have real impacts.

Design and development of data collection tools to support the development and processing of environmental and TEK is not trivial. As previous studies from the context of CSCW have demonstrated these usually rely heavily on collaborative tasks, or tasks which have the potential to bring the community closer together so that such knowledge can be effectively and accurately created (e.g. see Wulf et al., 2011; Pennington et al., 2007; Vitos et al., 2017). Considering that many of these communities are egalitarian, with cases where literally everyone in the community participates in the data collection and the development of community-generated TEK, make the relevance to the field of CSCW even more significant. Studies mainly emphasizing on the mapping interface, when this is used as the main interaction component to support this collaboration, also exist in the field of

participatory Geographical Information Systems (PPGIS) (e.g. see Brodnig and Mayor-Schonberger, 2000).

Moreover, with the majority of the communities mentioned above located in developing countries, issues such as technological infrastructure, familiarity with technology, education and literacy, local practices and environmental conditions are of utmost importance in terms of achieving seamless local human-computer interactions. Therefore, designing for these communities also traces back to the field HCI4D which is concerned with similar research questions to inform the development of information and communication technologies (ICT) in developing countries and where there is also a growing interest by the CSCW research community in terms of exploring how to *"bring new technology users from underserved communities into the fray"* (Kumar and Dell, 2018, p.2; Dillahunt et al., 2017). This exploratory paper contributes mainly to the discipline of ICT4D/HCI4D - and given the growing significance of TEK in this context – to the discipline of CSCW; we believe that with our empirical findings and methodological observations we will influence future work in both disciplines, especially with respect to designing for low-literate users.

Our emphasis is on interactions of low-literate users with mobile interfaces. It is expected that by 2025 mobile subscriptions will reach 5.9 billion, with growth mainly driven by developing (GSMA Intelligence, 2018). It therefore comes with no surprise that a growing research body explores mobile phone use (Dell and Kumar, 2016) - mainly basic or feature phones - and especially how low-literate users interact with them, as most mobile devices "assume a reasonable amount of literacy" (Dodson *et al*. 2013, p. 389). Currently, only a few studies examine how low-literate users interact with smartphones - despite their increasing lower costs and smartphone ownership being on the rise (Poushter *et al*. 2018). An even lower number of studies look into the design of interfaces that may assist low-literate users in data collection tasks which may further have the potential to support TEK in a collaborative context. In this paper, we build on methodological challenges discussed in the literature and examine the potential of carrying out an experiment with low-literate people in the UK to investigate the most successful interaction modes in a smartphone environment. Our experimental results will subsequently inform the interface design and additional experiments with end-users in Namibia and other regions in developing countries. We further reflect on our experience from carrying out experimental work for this type of participants in the UK, and we hope that our study will contribute to the evidence that it is being collected and which reports on how we can overcome some of the ICT4D methodological challenges by running usability studies with 'proxy' users in developed countries.

Background

UNESCO defines literacy as the ability of a person to read and write a short simple sentence in his or her everyday (UNESCO, 2006). Medhi et al. (2010) use the term low-literate to refer to: non-literate - i.e. those with an inability to read or write - and semi-literate - i.e. those that are able to read with difficulty. The authors suggest that low-literate people exceed the two billion worldwide. The term 'low-literate' in this study, as it is explained later in this paper, is used to refer to people with limited confidence in completing certain tasks, which assume a certain level of textual literacy in the developed world, and it further extends to include people with low digital literacy skills.

Early research on mobile phones for developing countries, uses ethnography to understand contextual characteristics and user needs (Chipchase, 2006; Belay an McCrickard, 2006; Dodson *et al*. 2013). Studies also carry out prototype development and usability evaluations to test mainly communication features (of basic phones, feature phones and occasionally smartphones) such us the phone's diary to make a call or the use of text-message functionality (Lalji and Good, 2008; Friscira *et al*. 2012; Dodson *et al*. 2013). Given a growing number of mobile phones are now connected to the Internet, research also explores the design of applications for water quality information and alerts (Brown *et al*. 2012); search for a job or navigating the city (Medhi *et al*. 2007) and health applications (Chaudry *et al*. 2012; Kumar and Anderson, 2015).

Although there are still a few studies which suggest augmenting rather than eliminating text-based features in ICTs for low-literate people (Knoche and Huang, 2012), a much higher number of research studies demonstrated that pictorial interfaces with little or no text are more useful (Parikh *et al*. 2003; Medhi *et al*. 2006; Medhi *et al*. 2007). Lack of education and literacy skills do not only influence one's ability to read text, but as Medhi *et al*. (2010) discuss, a person's cognitive abilities and linguistic sequential memory. One of the most notable implications of this is its direct effect in people's ability to understand abstractions, which are now commonly used in interface design and mainly for supporting hierarchical navigation and information structures. An increasing number of studies demonstrate low-literate people's difficulties in understanding and using menus that are based on hierarchies and instead recommend linear structures with up and down button or scrollbars to navigate them (Lalji and Good, 2008; Chaudry *et al*. 2012; Medhi *et al*. 2010; Winchiers-Theophillus *et al*. 2010]. It should be, however, noted that improved digital literacy and familiarity in terms of interacting with mobiles phones helps low-literate users overcome this problem and slowly develop similar proficiency levels in using their phones with those of literate users (Medhi *et al*. 2010).

Research further suggests that pictorial design should be fully embedded into cultural contexts, local meanings (Lalji and Good, 2008; Medhi *et al*. 2006) and

user preferences (Lalji and Good, 2008; Frommberger and Waidyanatha, 2017). There is evidence in the literature that low-literate users understand better hand-drawn, semi-abstracted graphics which incorporate action cues, while photo-realistic images are usually more effective in deeper interaction modes (Medhi *et al*. 2006). Additional modalities in the user interface such as audio feedback and voice annotation have been also tested and proved to be effective in specific contexts of use (Chipchase, 2006; Medhi *et al*. 2006; Deo *et al*. 2004; Medhi *et al*. 2007; Lalji and Good, 2008).

Previous research around input methods for basic or feature phones explores the use of keypad (Bailly *et al*. 2014; Lalji and Good, 2008) while few more recent studies investigating interactions with touchscreens (Chaudry *et al*. 2012; Friscira *et al*. 2012). Depending on the context of these studies and whether participants own a smartphone or not, there is consensus that low-literate users are hesitant with touching the screen of touchscreen devices and they are struggling with different types and outcomes of tapping. Friscira *et al*. (2012) suggest that low-literate participants should be first trained to the basics of smartphone touchscreen interaction. Despite these concerns, Chaudry *et al*. (2012) suggest the use of scrollbars on touchscreen, while Katre (2008) argued that low-literate users' lack of fine motor skills due to non-practice in writing makes thumb-based interaction more effective.

Although less popular compared to research around communication features, technologies (mainly PDAs and mobile phones) which are used to support low-literate users in data collection tasks have been around for some time (Vitos *et al*. 2013; Lewis and Nelson, 2006). Participatory mapping is a well-established methodology for obtaining knowledge from local communities concerning their living conditions and their environment. However, our focus here is on ICT technologies that could be used by the communities themselves, whereas in traditional participatory mapping exercises in this context, the documenting of resources and map-making was produced by expert cartographers with the communities' active assistance (Vitos, 2018). Examples from our context include: CyberTracker, a pictorial data collection interface, which has been used by non-literate trackers mainly in South Africa to support wildlife monitoring and natural resource management (Leibenberg *et al*. 2017); a smartphone-based app to collect georeferenced document and upload information that can support campaigns against illegal logging activities in Cambodia (Copenhagen Post, 2017); Extreme Citizen Science tool Sapelli, a pictorial smartphone-based interface which allows non-literate indigenous communities in Congo, Brazil, Cameroon, Namibia and others to collect any data that supports indigenous communities in knowledge co-production practices and which is used by non-literate (Vitos *et al*. 2017); the Sahana Disaster Management system that employs pictorial icons to check the emergency preparedness of low-literate communities in Philippines and provide

them with response and recovery information (Frommberger and Waidyanatha, 2017).

HCI research in this context is limited, with the majority of experiences remaining mostly anecdotal evidence; few of these experiences were presented in the Workshop on 'Lessons learned from volunteers' interactions with Geographic Citizen Science' which took place in London in April 2018 and which was organized by this paper's authors. The few existing findings are not different from the research discussed above. For example, Vitos *et al*. (2017) report that symbolic metaphorical conventions to represent categories in pictorial design do not work with low-literate people despite those being developed in participatory design workshops (**Figure 1**). Icons to represent specific objects had to incorporate action as they were taken too literally and therefore agree with (Medhi *et al*. 2006). Fear of using the technology and difficulties with the touchscreen, due to rough skin, or not understanding input methods (e.g. tapping and long clicks) have been also observed (Vitos *et al*. 2017; Vitos, 2018).

Figure 1:Community workshop for participatory pictorial design in Congo – Extreme Citizen Science project.

Evaluation of Sapelli, a data collection application which is based on a hierarchical navigation structure, is in line with previous research findings as low-literate people had difficulty understanding how to navigate it (Vitos *et al*. 2017). A physical interface was developed and evaluated to overcome Sapelli's challenges in the field; Tap&Map is a smartphone-based interface which uses near field communication (NFC) cards to tag an object together with each GPS coordinates (Vitos *et al*. 2017). Results demonstrated that participants had a 97.5% success rate in task completion using Tap&Map and they found it "faster, easier and more comfortable to use compared to Sapelli" (Vitos *et al*. 2017, p. 1584).

In this paper we consider research findings from the broader ICT4D field and previous work with Sapelli and Tap&Map to develop and further evaluate four user interfaces for data collection purposes with low-literate participants in the UK, which have the potential to further assist low-literate users in developing countries to perform collaborative tasks. For example, our findings will support the development of interfaces to collaboratively collect resource management data or data related to TEK, which is usually an important consideration with indigenous communities, so that the communities themselves can collaboratively identify solutions to local issues (e.g., wildlife crime in Cameroon, illegal logging in the Republic of Congo, resource management and fighting cattle invasions in Namibian Nyae Nyae Conservancy). It should be finally noted that although audio found to improve usability of otherwise problematic hierarchical structures (Vitos *et al*. 2017) we haven't explored this feature further, as it is not always an appropriate modality especially in high risk environments in terms of people's safety (e.g. when monitoring wildlife crime).

Aims and Study Design

Aims and Context

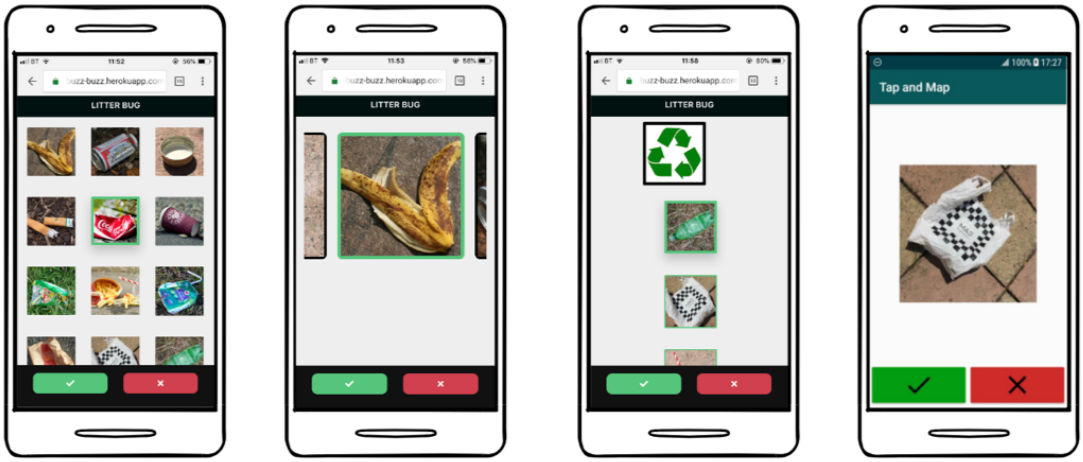
In this study we carried out a controlled experiment to evaluate four alternative user interfaces on a smartphone device, which have the potential to support low-literate users in data collection. Our goals are to evaluate: a. which interface is the easiest to use for the target user group and; b. which interfaces the users prefer to use.

One of the most widely recognised methodological implications in HCI4D research, is the difficulty in carrying out experimental work in remote locations, especially as part of an agile UCD approach. To make preliminary design choices which we could then test with users in developing regions we decided to explore how a representative user audience based in the UK, interacts with different interfaces. Within this context our first experimental design implication was to create a recognizable and meaningful task for our participants; a task preferably from the environmental context, which they could understand quickly, and which would involve the use of pictorial icons that they could immediately recognize and relate to them. Litter data collection is a task that we expected to appeal and be sensible to our participants and therefore it was the topic chosen for our experiment. Our research started with the design of initially 20 litter images (e.g. banana peels, cola cans, plastic carrier bags), which after a pilot study with five participants, were reduced down to 15 in order to remove unnecessary complexity which was overwhelming for our subjects and to further decrease the time required to run the experiment from six minutes to four minutes per task. Three images were also

deemed unclear during the pilot and therefore they were replaced, while the size of all images increased so they were easier to see upon recommendation of our pilot participants.

In the same pilot study we further tested that tasks and supporting materials were easy to understand. Although, we initially included a combination of icons and images to investigate user preference over different visualisations this combination found to be confusing. Despite the fact that previous research with low-literate users in developing countries suggests the use of hand-drawn, semi-abstracted images (Medhi et a., 2006) we decided to include only photo-realistic images, as we are aware of previous research in data collection with low-literate users in urban centers which suggests the use of photo-realism perhaps due to the fact that people in urban centres are more exposed to similar visual cues (Chiaravalloti, 2018).

The four interfaces that we evaluated in our study are shown in **Figure 2** and include: Icon Menu (Figure 2a); Swipe Menu (Figure 2b); Sapelli Menu (Figure 2c); Tap&Map (Figure 2d).



a. Icon menu b. Swipe menu c. Sapelli-inspired menu d. Tap&Map

Figure 2: Data Collection Interfaces tested in our study

The first two interfaces (i.e. Icon and Swipe menu) were designed to provide a provide a linear navigation structure (i.e. a structure which is not based in a hierarchy; it supports moving backward or forward in a sequence of objects) as previously suggested (Brumby and Zhuang, 2015; Cockburn *et al*. 2007; Lalji and Good, 2008; Chaudry *et al*. 2012; Medhi *et al*. 2010; Winchiers-Theophillus *et al*. 2010). The Icon menu showed a total of 12 images in two screens (i.e. three per each row) and required a single finger scrolling to navigate vertically between the two screens. The Swipe menu included the same images which were shown

horizontally, with one image shown per screen. Main interaction input was a horizontal single finger swiping (either left or right) to navigate across the images.

The Sapelli menu and Tap&Map interfaces were designed based on previous work on data collection with low-literate users (Vitos *et al*. 2017). Sapelli, provides a hierarchical menu structure which in our study had two top level categories for grouping the 15 litter items in recyclable and non-recyclable. Sapelli requires users to tap to select an image but it also requires them to correctly identify to which of the two categories the item falls and therefore navigate across this hierarchical menu structure. Tap&Map (Vitos *et al*. 2017) is a tangible interface and it requires very little interaction with the phone. The data items are shown on 15 physical NFC cards (i.e. one per each image). Participants browse the cards and once they identify the one they want to map, they tap the card against the phone and the image appears on the phone. Participants have to further confirm their selection by tapping the tick or cross icon on the phone (as shown on Figure 2d) which is the only interaction with the screen of this interface.

Experimental Design

Starting with a 15 minutes training session each participant was introduced to the basics of smartphone interaction, using each one of the four interfaces and the experiment's instructions and they were provided with either a Motorola Moto G or Samsung Galaxy Xcover 4 device, which they used to complete the tasks. The experiment required participants to complete a goal-oriented task using each of the four interfaces by matching the image on the interface with the appropriate litter type (i.e. total tasks n=4). To ensure that all litter types were equally used (rather than picking from those only physically present on a street), the 15 litter types were all shown as separate A4 paper printouts which were placed around participants before the experiment. Each task then required participants to map as many litter images as possible (out of total n=15) in 4 minutes using each of the four interfaces. A 'within subjects' design required all participants to complete the same tasks using all four interfaces and the interfaces were shown in a randomized order. Each experiment was carried by one of this paper's authors.

Task completion times and error rates were measured during the test using a timer (i.e., to measure the four minutes task duration) and observation notes (e.g.., an error occurred when a participant matched an icon to the incorrect A paper printout which was noted by the researcher observing the experiment). At the end of the experiment a score was calculated by summing each participant's number of correct matches and deducting the number of mistakes (i.e. Task Success = Total number of correct matches - Number mistakes). These scores were then averaged to provide an overall score for each interface. At the end of each one of the four tasks participants were verbally asked how they found the task, how confident they felt completing the task and how much they enjoyed this

version of the litter data collection application. These questions were consistently asked across all tasks and all participants to understand their subjective experience of using each interface. At the end of the experiment participants were again verbally asked which of the four interfaces they most and least enjoyed; the researcher conducting the experiments took notes of their answers which were then processed in the analysis. The experiments were also audio-recorded, data were transcribed and further analysed. Quotes from participants and research observations were amassed in addition to quantitative data to provide some qualitative insight. The data was used to produce a selection of pivot tables in Microsoft Excel to give a high-level overview of how each interface performed. This made it possible to detect trends and anomalies in the data. Individual quotes and observations were grouped into a number of themes that were analysed and turned into key findings.

Recruitment and Participants

Recruiting participants with low-literacy skills in the UK was a complex process. Within a period of over two months we contacted 50 organizations in the UK including adult learning centres, adult literacy learning groups, job centres, churches, community centres, local radio stations and so on. It is not uncommon for illiterate people to hide their lack of literacy and this is another obstacle HCI4D research which takes place in a western country has to overcome (Friscira *et al*. 2012; Knoche and Huang, 2012). To work around this problem we were slightly more flexible in terms of how the term 'low-literacy' is used in the HCI4D literature, to include people who were able to read or write a short message but with limited confidence in basic skills for life (as described by the UK Government 2011 Skills for Life Survey) and which assume a certain level of literacy (see also Kodagoda and Wong, 2008).

Prior to the experiment participants were asked questions to establish their age, gender, ethnicity, occupation, level of literacy and numeracy using the UK Government 2011 Skills for Life Survey (Department for Business, Innovation and Skills, 2012) and participants' confidence with technology using the Open University Digital Skills Checklist (The Open University, 2018).

Overall 13 participants took part in the experiment with an age range of 58-80 years old (avg=71; females=7; males=6); participants from this age group were less confident in their interactions with mobile phones (especially smartphones), which is usually a common characteristic in the indigenous communities we work with in remote areas. Two of the participants were completely illiterate; none of the participants were confident in using technology although 12 out of 13 owned a phone but mainly for phone calls and/or texting. Our low-literate participants (n=11) were confident writing a short physical message to friends and describing their medical symptoms to a physician but they were not confident withdrawing cash from an ATM cashpoint, reading a bus timetable and comparing products or

services. These tasks assume a certain level of literacy, which in some situations is taken as granted for completing every day tasks and in terms of interacting with digital technologies in the western world. A failure to show appropriate confidence levels and an ability to complete these tasks was a precondition for participant recruitment.

Results

As **Table 1** shows participants scored the lowest with Sapelli (TS=7.1), the highest when using the Icons menu (TS=10.7) followed by Tap&Map (TS=10.2) and the Swipe menu (TS=8.4). Participants commented on the usability of both the Icon menu (e.g. "I like seeing all the pictures together, that made it easy to use" - participant comment) and the Tap&Map interface (e.g. "…this was easy to use, the cards made it easy" - participant comment). Although Sapelli was used in this experiment with only two top level categories its hierarchical structure still confused participants. For example, one participant explained that "this was the hardest [interface to use] as you had to decide whether something was recyclable or not before finding it on the screen".

Table 1: Task Success [TS], Error Rates [ER], Standard Deviation [SD] and user preferences for each interface

	[TS]	[ER] (%)	Standard Deviation [SD]	Most Liked overall	Least Liked overall
Icon Menu	10.7	9.4	4.8	2	1
Swipe Menu	8.4	4.2	4.2	0	8
Sapelli	7.1	2.5	3.09	1	3
Tap&Map	10.2	2	3.3	10	1

Although our population sample is small to draw any concrete conclusions to link results to user demographics, we further observed that participants who had no prior experience in using a smartphone performed better using Tap&Map (TS=8.0), followed by the Icon Menu (TS=6.6), the Swipe menu (TS=4.8) and finally Sapelli (TS=4.0).

Although the Icon menu scored the highest in terms of task success, it was also the interface where we observed most errors taking place. However, participants managed to recover easily from their errors and hence complete their tasks successfully. We believe that it was the interface's usability that paradoxically led users to make more errors since the observer noticed that users became

overconfident and rushing through the task when using the Icon menu. In terms of error rates Tap&Map was the most successful, with a 2% error rate, followed by Sapelli (ER=2.5%) and the Swipe menu. (ER=4.2%).

Ten of the thirteen participants liked the Tap&Map interface the most. The one participant who disliked Tap&Map had arthritis, which caused a lack of dexterity in his hands and therefore difficulty in handling the NFC cards. Interestingly enough the least liked interface was the Swipe menu; during the experiment participants observed to struggle with the one finger swiping interaction which caused frustration to some (e.g. "this one [Swipe menu] didn't adapt to me, it wasn't easy and it was quite frustrating" – participant comment).

We also asked participants at the end of each task to rate each interface in terms of its perceived usability, how confident they felt completing the task and how much they enjoyed using it using a four-likert scale. Tap&Map and the Icon menu scored the highest in terms of perceived usability and confidence, followed by Sapelli and the Swipe menu. Participant 13 who was illiterate and had never held a smartphone before commented about Tap&Map "I could do that all day, I am used to not being able to do anything on a phone, maybe I am not that thick after all…this gives me a lot of confidence that I am not as thick as I think I am".

At the end of the experiment, once participants had experienced all four interfaces, they were asked which interface they most and least liked using. The Swipe menu was the interface our participants liked using the least (8/13) while Tap&Map was the interface participants liked using the most (10/13). Three participants did not particlarly enjoy using Tap&Map, with two participants preferring to use the Icon menu (2/13). These three participants who did not enjoy using Tap&Map experienced some physical discomfort while using Tap&Map which was not surprising due to the age of the participants – and led to their lack of enjoyment.

Discussion and conclusions

Building strong sustainability agendas which have the potential to truly impact and transform our world, amongst others, requires zooming into local environments and providing the mechanisms that let people look into issues they face locally, and supporting them in the identification of effective solutions to address them. Data collection tools are becoming increasingly popular in terms of supporting users with these endeavors. Low-literacy and the limited prior experience of users in interacting with technological artefacts need to be taken into account when designing for these particular user audiences. Taking into account existing HCI4D literature in this study we developed and evaluated four alternative interfaces to support low-literate users in data collection tasks using smartphone devices.

Building on research we suggests that a linear navigation structure works well with low-literate users, we developed two interfaces which had a linear navigation

but required different types of interaction. We found that a linear navigation, which involves minimum interaction with the smartphone, was the most successful interface (i.e. Icon Menu) in this study, for its usability and the second most preferred by our participants. We also observed that a linear structure can result in a very negative user experience and reduce usability when users are expected to constantly interact with their device, as it was the case with the Swipe menu, which was although achieved a higher score success rate than Sapelli it was the least like interface. It should be noted that there are no other studies to report a negative user experience associated with linear navigation structures and therefore this finding might need further investigation in other contexts of interacting with smartphone devices.

The second most successful interface in terms of task success was Tap&Map, which scored the highest for user preference. Tangible interfaces have the potential to keep interaction with the phone at its minimum and this was appreciated by the majority of our participants. It should not be, however, noted that the average age of our participants was 70.9 years old and some of them suffered from arthritis, which caused difficulty navigating across the pack of the NFC cards, therefore we suggest that further testing is required to assess how usable are tangible interfaces for data collection purposes in various environmental conditions and for various user groups.

There is already evidence in previous research both from the context of data collection but also mobile phone use in a broader sense, that hierarchical structures are problematic with low-literate users, and our results agree with those findings. Although Sapelli scored the second lowest error rate, it was still the least usable interface in terms of its task success rate. This highlights the importance of another usability principle, error recovery, which is much more problematic in hierarchical navigation structures since users once they get into the lowest levels of a decision tree find it harder to go back and recover from any errors compared to recovering from an error in a linear navigation structure (e.g. the Icon menu had the highest error rate, yet it was the most usable interface in terms of task success). From this finding we suggest that future research related to hierarchical navigation structures should look into error recovery and interface design cues that have the potential to release users from the already increased cognitive effort that hierarchical structures require. Such features could have a significant impact when a hierarchical navigation structure is the only option.

HCI4D researchers explain how conducting HCI research in developing countries has unique challenges due to sociocultural, linguistic and other implications (Anokwa *et al*. 2019; Chetty and Grinter, 2007). One major obstacle to implementing a user-centred design approach to support the development of extreme citizen science data collection tools is proximity and access to the target user audiences. In other words, constant development and evaluation of prototypes with target users in the field is not always feasible neither it is possible to carry out

complex experimental designs which rely on evaluating a high number of prototypes in one go. At the same time running usability studies in a western country which require the recruitment of low-literate (or even proxy) users has its own challenges and still results may be biased as they are open to influences from local socio-cultural and environmental conditions which are significantly different from those in the field. In an attempt to deal with all these challenges, we tested four prototypes with a relevant participant audience in the UK. Although it may be argued that this study's population sample is small to run an in-depth analysis and provide concrete conclusions, it has provided us with enough insight in terms of choosing the two most successful interfaces which we then tested with low-literate users in Namibia for collecting data for natural resource management purposes. Our preliminary findings from the field testing agree with the usability study that we describe in this paper. To further evaluate the validity of this approach, we are planning to incorporate more testing of our interfaces and tools in developed countries with representative user audiences, in preparation of and prior to testing in remote locations, as others have also recommended (e.g.Chetty and Grinter, 2007; Knoche and Huang, 2012). We believe that providing evidence and reflecting on the results and effectiveness of these experimental approaches may significantly help tackle some of the most critical methodological challenges in HCI4D research.

ACKNOWLEDGMENTS

We would like to thank the Extreme Citizen Science research team for their input in this study's experimental design and the people who participated in our study. This research project is funded by European Research Council's project Extreme Citizen Science: Analysis and Visualisation under Grant Agreement No 694767.

REFERENCES

Anokwa, Y. et al. (2019). 'Stories from the field: Reflections on HCI4D Experiences'. *Information Technologies and International Development*, vol. 5, pp. 101-115.

Bailly, G., Oulasvirta, A., Brumby, D., Howes, A. (2014). Model of visual search and selection time in linear menus. In Proceedings of the SIGCHI Conference on human factors in computing systems, (Toronto, Canada), pp. 3865-3874.

Belay, E. and McCrickard, S. (2006). 'Mobile User Interaction Development for Low-Literacy Trends and Recurrent Design Problems: A Perspective from Designers in Developing Country'. *Mobile User Interaction Development*, pp. 409-417.

Brodnig, G., Mayer-Schönberger, V. (2000). Bridging the Gap: The Role of Spatial Information Technologies in the Integration of Traditional Environmental Knowledge and Western Science. The Electronic Journal of Information Systems in Developing Countries, vo. 1, no. 1, p. 1-15.

Brown, D., Marsden, G., Rivett, U. (2012). Water Alert!: Using Mobile phones to Improve Community Perspective on Drinking Water Quality in South Africa, In Proceedings of the 5th International Conference of Information and Communication Technologies and Development, Atlanta, USA, pp. 230-240.

Brumby, D., Zhuang, S. (2015). Visual grouping in menu interfaces. In Proceedings of the 33rd Annual ACM Conference on human factors in computing systems, Seoul, Korea, pp. 4203-4206.

Chaudry, B., Connelly, K., Siek, K.,Welch, J. (2012). Mobile interface design for low-literacy populations. In Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium: pp. 91-100.

Chetty, M., Grinter, R. (2007). HCI4D: HCI challenges in the global south. In CHI '07 extended abstracts on Human factors in computing systems, San Jose, USA, pp. 2327–2332.

Chiaravalloti, R. (2017). Local communities and conservation in the Pantanal wetland, Brazil. *Doctoral dissertation. University College London*, London, UK.

Chipchase, J. (2006). 'How Do You Manage Your Contacts If You Can't Read or Write? Interactions'. *Interactions,* vol. 13, no. 6, pp. 16-17.

Cockburn, A., Gutwin, C., Greenberg, S. (2007). A predictive model of menu performance. In Proceedings of the SIGCHI Conference on human factors in computing systems, San Jose, USA, pp. 627-636.

Copenhagen Post. (2017). Danish app helping Cambodian activists to combat illegal deforestation. http://cphpost.dk/news/danish-app-helping-cambodian-activists-to-combat-illegal-deforestation.html (Accessed 15 April 2018).

Dell, N., Kumar, N. (2016). The Ins and Outs of HCI for Development. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, USA, pp. 2220-2232.

Deo, S., Nichols, D., Cunningham, S., Witten, I., Trujillo, M. (2004). Digital library access for illiterate users. In Proceedings 2004 International Research Conference on Innovations in Information Technology (IIT2004,) Dubai, pp. 506–516.

Department for Business, Innovation and Skills. (2012). The 2011 Skills for Life Survey: A Survey of Literacy, Numeracy and ICT Levels in England. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/36000/12-p168-2011-skills-for-life-survey.pdf#page=66 (Accessed 1 June 2018).

Dodson, L., Sterling, R., Bennett, J. (2013). Minding the Gaps: Cultural, Technical and Gender-based barriers to mobile use in oral-language Berber communities in Morocco. In Proceedings of the Sixth International Conference on Information and Communication Technologies and Development, Cape Town, South Africa, pp. 79-88.

Dillahunt, T., Erete, S., Galusca, R., Israni, A., Nacu, D., Sengers, P. (2017). Reflections on Design Methods for Underserved Communities. CSCW '17 Companion, February 25 - March 01, 2017, Portland, OR, USA ACM 978-1-4503-4688-7/17/02.

Friscira, E., Knoche, H., Huang, J. (2012). Getting in touch with text: Designing a mobile phone application for illiterate users to harness SMS. In Proceedings of the 2nd ACM Symposium on Computing for Development, Atlanta, USA.

SynchronicityEarth.org, 2018. Embracing biological and cultural diversity: An Interview with Dr Jerome Lewis. Synchronicity Earth. (2018). https://www.synchronicityearth.org/embracing-biological-and-cultural-diversity-an-interview-with-dr-jerome-lewis/ (Accessed 2018-09-11).

Frommberger, L. and Waidyanatha, N. (2017). Pictographs in Disaster Communication for Linguistically Challenged and Illiterate Populations: A Survey on Background and Existing Practices. *International Journal of Information Systems for Crisis Response and Management*, vol. 9, no. 2, pp.37-57.

Katre, D. (2008). One-handed thumb use on smart phones by semi-literate and illiterate users in India: A usability report with design improvements for precision and ease. In Workshop on Cultural Usability and Human Work Interaction Design - NordiCHI Conference, Lund, Sweden, 2008.

Kodagoda, N. and Wong, W. (2008). Effects of Low and High Literacy on User Performance in Information Search and Retrieval. In Proceedings BCS-CHI-08 People and Computers: Culture, Creativity, Interaction - Volume 1, pp. 173-181.

Knoche, H., Huang, J. (2012). Text is not the enemy: How illiterates' use their mobile phones. In Proceedings of the CHI Conference on Human Factors in Computing Systems, Auston, USA.

Kumar, N., Anderson, R. (2015). Mobile Phones for Maternal Health in Rural India. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, April 18-23, Seoul. Korea, pp. 427-436.

Kumar, N and Dell, N. (2018). Towards Informed Practice in HCI for Development. In Proceedings of the ACM on Human-Computer Interaction, Vol. 2, CSCW, Article 99 (November 2018). ACM, New York, NY. 1-20.

Lalji, Z., Good. J. (2008). 'Designing new technologies for illiterate populations: A study in mobile phone interface design'. *Interacting with Computers,* vol. 20, no. 6, pp. 574-586.

Leibenberg, L., Steventon, J., Brahman, !N., Benadie, K., Minye, J., Langwane, H., Xhukwe, Q. (2017). 'Smartphone Icon User Interface design for non-literate trackers and its implications for an inclusive citizen science'. *Biological Conversation*, vol. 208, pp. 155-162.

Lewis, J., Teodyl, N. (2012). Accessible technologies and FPIC: independent monitoring with forest communities in Cameroon. In Participatory Learning and Action 65. Ed. by Krystyna Swiderska, Kanchi Kohli, Harry Jonas, Holly Shrumm, Wim Hiemstra and Maria Julia Oliva. Participatory Learning and Action 65. IIED. Chap. 13, pp. 151–165. ISBN: 978-1-84369-851-7

Lewis, J., Nelson, J. (2006). Logging in the Congo Basin: What Hope for Indigenous Peoples' Resources and their Environments? In International Work Group for Indigenous Affairs (IWGIA). Vol. 4, no. 06, pp. 8–15.

Medhi, I., Sagar A., Toyama K. (2006). 'Text-Free User Interfaces for Illiterate and Semi-Literate Users'. *Information and Communication Technologies and Development*, vol. 4, no. 1, pp. 37-50.

Medhi, I., Prasad, A. and Toyama K. (2007). Optimal audio- visual representations for illiterate users. In Proceedings of the 16th international conference on World Wide, (Alberta, Canada), pp. 873-882.

Medhi, I., Gautama, S., Toyama, K. (2009). A comparison of mobile money-transfer UIs for non-literate and semi-literate users. In Proceedings of the SIGCHI Conference on human factors in computing systems, (Boston, USA), pp. 1741-1750.

Medhi, I., Raghu Menon, S., Cutrell, E.,Toyama, K. (2010.) Beyond Strict Illiteracy: Abstracted learning among low literacy. In Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development, (London, UK), 1-9.

Parikh, T., Ghosh K., Chavan, A. (2003). Design Considerations for a Financial Management System for Rural, Semi-literate Users. In Proceedings of Human Factors in computing Systems, (Florida, USA), pp. 824-825.

Pennington, D.D., Madin, J., Villa, F., & Athanasiadis, I.N. (2007). Computer-supported collaborative knowledge modeling in ecology. Workshop on social and collaborative construc- tion of structured knowledge. In *16th International World Wide Web Conference (WWW2007). CEUR Workshop Proceedings, 273*. BANFF, Canada published online.

Poushter, J., Bishop, C., Chwe, H. (2018). Smartphone ownership on the rise in emerging economies. http://www.pewglobal.org/2018/06/19/2-smartphone-ownership-on-the-rise-in-emerging-economies/ (Accessed 09 Sept. 2018).

GSMA Intelligence. The Mobile Economy: Sub Saharan Africa (2018). https://www.gsma.com/mobileeconomy/sub-saharan-africa-2017/ (Accessed 30 March 2018).

The Open University, 2018. Being digital: Digital literacy skills checklist. Available from: http://www.open.ac.uk/libraryservices/beingdigital/accessible/accessible-pdf-35-self-assessment-checklist.pdf (Accessed June 2018).

UNESCO (2006). Chapter 6 understandings of Literacy. http://www.unesco.org/education/GMR2006/full/chapt6_eng.pdf (Accessed 9 Sept. 2018).

Vitos, M. (2018). Making local knowledge matter: Exploring the appropriateness of pictorial decision trees as interaction style for non-literate communities to capture their traditional ecological knowledge. *Doctoral dissertation*. *University College London*, London, UK.

Vitos, M., Stevens, M., Lewis, J., Haklay, M. (2013). Making local knowledge matter. In Proceedings of the 3rd ACM Symposium on Computing for Development - ACM DEV '13. New York, New York, p. 1.

Vitos, M., Altenbuchner, J., Stevens, M., Conquest, G., Lewis, J., Haklay, M. (2017). Supporting Collaboration with Non-Literate Forest Communities in the Congo-Basin. In CSCW '17 Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. (New York, USA), pp. 1576-1590.

Winchiers-Theophillus, H., Bidwell, N., Blake, E., Kapuire, G., Rehm, M. (2010). Merging experiences and perspectives in the complexity of cross-cultural design. In Proceedings of the 9th International Workshop on Internationalization of Products and Systems, London, UK, pp. 131-140.

Wulf, V., Rohde, M., Pipek, V., Stevens, G. (2011). Engaging with Practices: Design Case Studies as a Research Framework in CSCW. CSCW 2011, March 19 23, 2011, Hangzhou, China.

Schwarzer, J.; von Luck, K.; Draheim, S.; Koch, M. (2019): Towards Methodological Guidance for Longitudinal Ambient Display In Situ Research. In: Proceedings of the 17th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing an the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep07

Towards Methodological Guidance for Longitudinal Ambient Display In Situ Research

Jan Schwarzer[†], Kai von Luck[†], Susanne Draheim[†], Michael Koch[‡]
[†]Hamburg University of Applied Sciences
{jan.schwarzer, kai.vonluck, susanne.draheim}@haw-hamburg.de
[‡]Universität der Bundeswehr München
michael.koch@unibw.de

Abstract. Field deployment research represents a promising way for understanding how technology is utilised in the wild. It gained relevance in both HCI and CSCW, and allows, for instance, to investigate how technology is socially embedded in real world contexts. However, such enterprises are considered complex in nature due to continuously changing conditions such as practices surrounding technology. In situ research has yet to gain momentum, leaving researchers with little theoretical guidance. In response, the present paper proposes the application of classic grounded theory in longitudinal field deployment studies for ambient displays. We argue that the methodology is a valuable choice in coping with the challenges surrounding in situ evaluations and simultaneously ensures methodological rigour. This paper contributes a practical systematisation of the methodology's two core concepts, namely constant comparison and theoretical sampling. It sheds light on their exemplary application in investigating quantitative interaction data in the early stages of our ongoing research. With that, we hope to encourage future research and provide a first stepping stone towards developing methodological guidance for evaluations of ambient displays in the wild.

Introduction

Socially embedded technology arguably challenged and changed practices like no other technological artefact had done before (Wulf et al., 2015). Research in this context embarks on affecting the fundamental ways of how research is conceptualised, as people and practices are more than just their relationships with technology—the concept of the 'user' becomes problematic (Bjørn and Boulus-Rødje, 2015). Here, practices and technology are being understood as intertwined and as continuously changing entities (Bjørn and Boulus-Rødje, 2015). Despite its challenging nature, research in the wild sheds light on real usage and highlights how technology interacts with environmental conditions such as technology already in use (Siek et al., 2014). Contrary to other research approaches, field deployment research is capable of investigating longitudinal effects and enabling researchers to cope with issues such as the novelty effect (Alt et al., 2012). Generally, long-term deployment studies are considered promising as they allow one to scrutinise how a technology is adapted in a particular context (Alt et al., 2012; Preim et al., 2018; Siek et al., 2014). In fact, operational feasibility of novel technology can only be determined in the field (Nunamaker Jr. et al., 2015). Although, field deployment research is considered a messy enterprise (Alt et al., 2012; Siek et al., 2014), it simultaneously affects most notably both science and society (Nunamaker Jr. et al., 2015). In situ research has gained momentum throughout HCI and particular relevance in the CSCW and Ubicomp disciplines (Siek et al., 2014) as their data can be assumed of high value (Alt et al., 2012). Recent contributions from a variety of domains such as information systems (Nunamaker Jr. et al., 2015), HCI (Börner et al., 2013; Hazlewood et al., 2011; Matthews et al., 2007; Messeter and Molenaar, 2012; Siek et al., 2014), information visualisation (Preim et al., 2018), and CSCW (Bjørn and Boulus-Rødje, 2015) stress the relevance of in situ research. Some authors place a particular emphasis on calls for long-term in the wild evaluations (Börner et al., 2013; Hazlewood et al., 2011; Preim et al., 2018).

In ambient display research, a strong technology-driven focus is observable, failing to consider how actual operation relates to people's everyday lives (Matthews et al., 2007). Social aspects, per se, have received little attention and broadening the scope of investigation is recommended (Messeter and Molenaar, 2012). Evidently, real-world evaluations indicate a lack of methodological guidance as traditional evaluation approaches do not apply. For instance, Hazlewood et al. (2011) conclude that ambient displays, in all their forms, require methodological development for in the wild evaluations.

In response, this paper introduces a holistic methodological approach applied in our ongoing longitudinal in situ evaluation of our custom 'Ambient Surfaces' solution. The paper sheds light on the utilisation of classic grounded theory (GT) and the practical systematisation of its two core processes, namely constant comparison and theoretical sampling (Boeije, 2002). To this end, their procedure is exemplary illustrated by discussing quantitative analyses of interaction data during

the early stages of our research. By following classic GT methodology, we argue that we can both adhere to methodological rigour and maintain flexibility in light of the aforementioned challenges.

The paper is organised as follows: firstly, an overview of the current state in evaluating socially embedded technology in the wild is presented, primarily focusing on the domain of ambient displays. Secondly, the methodological foundations of our research are highlighted, including the research setting and purpose, a consideration of GT's fit to conduct longitudinal in situ research, and an introduction to GT methodology. Thirdly and foremost, the paper presents the practical systematisation of the two core processes. Finally, before concluding the paper with some recommendations for future research, it discusses our approach, presents contributions to existing knowledge, and illustrates some limitations.

Evaluation of Ambient Displays in the Wild

Discussions in the HCI literature on how to conduct evaluations of user interfaces range back to early work such as Bannon and O'Malley (1984). However, these discussions often either focus on laboratory settings or short-term in situ deployments. Siek et al. (2014) provide some guidance on how to realise field deployments in HCI, including how to design data collection instruments. However, they stay rather superficial on this issue.

Some information on the utilised methods for evaluating long-term deployments of ambient displays can be found in reports about single research projects such as Peltonen et al. (2008), Rogers et al. (2010), Alt et al. (2012), Ojala et al. (2012), Memarovic et al. (2016), and Shelton and Nesbitt (2017). However, while addressing the topic of gathering and analysing data, the authors usually do not provide any background motivation as to why they conducted their evaluations in the way it was demonstrated.

Börner et al. (2013) report that reviewed studies used a variety of methodologies. However, it seems that the literature review rather lists a set of different data collection methods (e.g. interviews and observations) and design approaches (e.g. user-centred design and exploratory design). Following Crotty (1998), we found that surveys and heuristic inquiries were the only identified methodological choices.

Input regarding potential research goals can be found in the work from Nunamaker Jr. et al. (2015). The authors address the broader field of information systems and structure research into the three phases: 'proof-of-concept' research, 'proof-of-value' research, and 'proof-of-use' research. For each phase, the authors envision both field studies and laboratory studies as valid approaches and list potential evaluation goals. They argue that "The research is not complete ... until proof-of-use research demonstrates that a self-sustaining and growing community of practice has emerged around the solution" (Nunamaker Jr. et al., 2015, p. 43). However, their contribution fails to provide further guidance on how to conduct the respective evaluations.

In conclusion, there is currently no methodological advice on how to conduct research in longitudinal ambient display deployments.

Methodological Foundations for Longitudinal Ambient Display In Situ Research

This paper aims to provide a first stepping stone towards developing methodological foundations for long-term ambient display research in the wild. This development process builds on an ongoing study, where we utilise our custom Ambient Surfaces solution. The study seeks to understand, how the solution is appropriated in an authentic environment. We propose building this knowledge-seeking process on classic GT as *a* way of conducting such enterprises. To this end, this paper illustrates a practical systematisation of GT's two core processes—constant comparison and theoretical sampling. While the next section thoroughly elaborates this practical systematisation, the following emphasises the overall research setting and purpose, envisions classic GT as a means to cope with the challenges in field deployment research, and briefly introduces GT methodology.

The Study: Research Setting and Purpose

For field deployment research, the complexity and scope are embodied in choices made with respect to the *target population*, *scale*, and *duration* (Siek et al., 2014). Accordingly, this paper subsequently focuses on these topics. More details regarding the research setting and the custom solution can be found in previous work (Schwarzer et al., 2016).

The field deployment commenced in February 2014 with one large and interactive screen, while a second one followed in August 2015. Data collection is still ongoing in 2019. Our Ambient Surfaces solution is deployed in a German company that can be characterised as a large-scale agile software development environment with eight agile teams at present (Dingsøyr et al., 2014). Foremost, the practice of Scrum (Schwaber and Beedle, 2001) is adhered to and accompanied by selected practices from Extreme Programming (Beck, 2000). The Ambient Surfaces were located in two distinct locations and show information from different tools utilised in the department. In the first setting, roughly 70 to 80 people had access to the screens—this included Scrum Masters, Product Owners, management personnel, and foremost software developers (i.e. almost two-thirds). Around 90% of these staff members were between 31 and 50 years old and approximately 75% of them had been working in the company for at least three years. The number of potential users increased substantially to over 400 from the entire company's campus in the second setting (including further management personnel and consultants). This is due to the fact that both systems were relocated in 2017 to a newly constructed building which also includes a canteen. Generally, a large number of passers-by is characteristic for this new setup as it was for the old one. For in-

Figure 1. Both Ambient Surfaces in their current setup as of 2019, situated in a hallway near the canteen. In this instance, the left system mostly shows contents from the 'Confluence View' (e.g. architecture articles) and the right screen illustrates information from the 'Test Suites View' (e.g. list of test suites).

stance, people typically have to walk past the systems when having lunch in the canteen. Figure 1 shows the setup in its current configuration as of 2019.

Fundamentally, the study sets out to contribute missing longitudinal findings of ambient displays in real world contexts (Schwarzer et al., 2016). Our ongoing multiple-year enterprise embarks on generating a substantive theory—i.e. a theory that sheds light on a particular empirical area in the real world (Glaser, 1978)—which conceptually explains how the Ambient Surfaces solution is appropriated in this particular setting. Foremost, we are interested in what the literature refers to as 'naturalistic usage' (Siek et al., 2014) rather than, for instance, usage originating in instances of a novelty effect (Koch et al., 2018). We position our research in the domain of proof-of-use research as we are largely dealing with issues surrounding operational feasibility (Nunamaker Jr. et al., 2015).

Fit of Classic Grounded Theory

To rigorously strive towards our research goal, we pondered over an appropriate research methodology. Due to the issues below, we finally decided to utilise classic GT as the methodological foundation for our research:

1. Fundamentally, GT methodology sheds light on *social* phenomena, independently of a particular research discipline (Glaser and Strauss, 1967). It therefore assists in coping with the issue that socially embedded technology cannot be investigated without its social components (Bjørn and Boulus-Rødje, 2015).

2. Furthermore, GT enables one to deal with the dynamic nature of field deployments as it is considered messy and may require changes in the data collection procedure (Siek et al., 2014). GT does not ask for preconceiving of any sort of data but asks to let the data emerge and to openly chose the most appropriate data collection method (Glaser, 2008).

3. Classic GT utilises both quantitative and qualitative methods beyond boundaries of specific research paradigms such as positivism and constructivism as it is considered a general methodology (Glaser, 1998). In fact, Glaser (1998, 2008) considers all kinds of data as valuable in the process of generating theory (e.g. documents, magazines, and interviews). In comparison to other GT variants such as Straussian GT (Corbin and Strauss, 2015), classic GT suits the requirements of field deployment research arguably better as, typically, a mixture of different methods (e.g. interviews, observations, and log files) is utilised in such endeavours (Alt et al., 2012). This methodological openness primarily led to the decision to commence our research with classic GT.

4. As long-term in situ research is such an unexplored territory (Börner et al., 2013; Hazlewood et al., 2011; Preim et al., 2018), scarce theoretical guidance arguably exists to pose any initial research questions or hypotheses. In fact, ambient display research lacks existing general theories (Alt et al., 2012). GT follows the notion of starting any research open-minded without any preconceived problem statements (Glaser, 2008). It asked the open question of "What's going on[here?]" (Glaser and Strauss, 1967, p. 97), which is—in one form or the other (e.g. Glaser and Strauss, 1967; Glaser, 1978, 1992)—the opening question in every GT study (Charmaz, 2006).

5. Proof-of-use research generally faces the issue of externalising and codifying a researcher's tacit knowledge (Nunamaker Jr. et al., 2015). GT provides a means to report a researcher's own experiences and thereby increases the traceability and credibility of a study (Boeije, 2002). Fundamentally, GT strives towards situating study participants' actions and interpretations in the relevant circumstances and thus making them explicit (Morse et al., 2009).

Introduction to Grounded Theory Methodology

In the 1960s, GT was an inductive response to predominant hypothetico-deductive research approaches (Glaser and Strauss, 1967; Morse et al., 2009; Stol et al., 2016). Over the last decades, however, GT evolved from its origins into two major streams: since the 1990s referred to as 'Glaserian GT' or 'classic GT' (Glaser, 1978, 1992, 2006; Stern, 1995) and 'Straussian GT' (Corbin and Strauss, 2008, 2015; Strauss, 1987; Strauss and Corbin, 1990, 1998). Following this development, further variants emerged, which are summarised under the term 'second generation' (Morse et al., 2009; Muller, 2014). According to Morse et al. (2009), differences in GT variants arise from epistemological stances, methodological

strategies, assumptions about what constitutes theory, and lastly conceptional directions.

Principally, GT represents a way of thinking about and conceptualising based on data (Morse et al., 2009). It is aimed at proposing grounded hypotheses, not facts (Glaser, 1978, 2008). Descriptions are put forth to foster an understanding of the rationale behind hypotheses (Glaser, 1978). Generally, it is geared towards discovering *a* not *the* theory (Heath and Cowley, 2004). Following Glaser (1978, p. 4), "... a theory should be able to explain what happened, predict what will happen and interpret what is happening in an area of substantive or formal inquiry." The process of generating theory is a continuous one of modification. Glaser (1978) refers to this attribute of a theory as 'modifiability'. GT is acknowledged to be a methodology which is a highly individually conducted endeavour (Morse et al., 2009). It fundamentally turns a human weakness into a strength as it allows the researcher to theorise about data during analyses whilst explaining data to oneself or colleagues (Muller and Kogan, 2012). As Muller and Kogan (2012) further note, GT seeks to formalise this cognitive process into a quality process to generate new insights and theories.

Practical Systematisation of the Analysis Process

Guided by other GT studies that structured their research in different phases (e.g. Boeije, 2002; Walsh, 2015), we organised our research in *four* research phases. In the first phase, only quantitative interaction data was considered. Subsequently, observations, a group interview, and a survey enriched the theory generating process in the second phase. While during the third phase statistical tests were additionally conducted, the last stage is ongoing and incorporates a respondent validation to conclude the research.

Below, it is concentrated on the practical systematisation of GT's two core processes during the first stage, denoted as 'Phase 1'. Therefore, we initially describe the theoretical underpinnings of the analysis process and subsequently highlight findings obtained by adhering to this rationale.

How the Analysis was Conducted

Fundamentally, our work builds on a constant comparison step-by-step approach introduced by Boeije (2002). She suggests *four* distinct criteria to be elaborated throughout each step: firstly, the analysis activities (i.e. a description); secondly, the aim of comparisons; thirdly, important questions asked and, lastly, the findings (see next section). In the following, it is now continued with the application of the first three of these four criteria in Phase 1, starting with the aim of comparisons, the questions, and finally a description of the comparison process.

Aim of Comparisons

The primary aim of comparisons in Phase 1 was to reveal latent patterns in interaction data. To this end, the identified patterns were used to state initial hypotheses about the actual utilisation, to pose new questions, and consequently to guide future data collection activities. It was intended to generate descriptive figures which provided a first theoretical glance at the issue of utilisation.

Important Questions

The following two questions were of particular relevance in Phase 1:

1. *During what daytimes is the Ambient Surface most prominently being utilised?* This question aimed at shedding light on the first emerging latent patterns that were prevalent in the material.

2. *How long can a novelty effect be notably observed in the collected data?* With this question, it was intended to further investigate the anticipated novelty effect at the beginning of the study.

Description of Comparison Activities

In particular, the guidelines regarding the use of quantitative data in GT had been consulted in Phase 1 (Glaser, 2008). Accordingly, the concept of *crude indices* had been applied, mainly for two reasons. Firstly, they reportedly "... suffice to indicate the concepts of the theory ..." (Glaser, 2008, p. 41). Secondly, the material at hand felt to be unsuitable to proceed any further with respect to the elaboration analysis introduced in quantitative GT as item discovery was not the issue at hand.

Below, the items of the crude index (i.e. the comparison candidates), the memoing process, and the sampling strategy are introduced.

Declaring Comparison Candidates The relevant interaction data log file included different variables containing information relating to *touch events*, which are triggered in the software framework when a person interacts with the display's surface (e.g. variable *Timestamp of Event*). Analogously to Glaser (2008), every variable in this file represented an *item* for a possible crude index. The general aim is to create a crude index, incorporating at least two items to indicate the concepts of a theory by leveraging cross-tabulations. However, the selection of items is a highly individual choice. Glaser (2008) refers to this process as a type of piloting study, given that the researcher literally plays with a set of different items.

The variable *Timestamp of Event* seemed most promising, primarily due to its nature to describe usage over time. A crude index *Utilisation of the Ambient Surface* consisting of two items was created. Firstly, the item *Number of Interactions*, which summarises data from *Timestamp of Event*. Secondly, the item *Daytimes* was utilised. This variable cuts *Number of Interactions* in temporal segments (i.e.

27 half-hourly segments from 07:00 to 20:00). The chosen time frame felt reasonable as it happened to account for the majority of interactions.

Memos and Diagrams Fundamentally, Glaser (1978) considers memos as the core stage of each GT research endeavour. Yet, he fails to clearly explain how memos and diagrams were utilised during the elaboration analysis in quantitative GT (Glaser, 2008). However, other books on GT provide sufficient detail on this matter (e.g. Charmaz, 2006; Corbin and Strauss, 2015; Glaser, 1978, 1998).

One distinct feature of memos and diagrams in this work is that they are primarily digital and, in some cases, printed pieces of data artefacts. Corbin and Strauss (2015), on the contrary, utilised long hand-written memos to reflect on their research process. While they found that qualitative data entails complex and cumulative thinking, this work initially considered only quantitative data. However, Corbin and Strauss (2015) acknowledge that memo-writing is a highly individual process and conclude that the important part is that memos are created, especially in longitudinal research.

A Microsoft Excel file was created to store memos, incorporating the aforesaid crude index with its items in cross-tabulations. This file also left room for additional notes and comments. For example, software bugs, change requests, and relevant correspondences were also documented in this file. A memo wall in the authors' office space was utilised to collaboratively reflect on data and simultaneously to inspect and sort memos at a glance (see Figure 2).

Figure 2. Parts of the utilised memo wall displaying a diverse set of touch events statistics, user interface screenshots, and charts in the authors' office space.

Theoretical Sampling The following suggestions were considered in approaching the sampling procedure: firstly, Muller (2014) encourages researchers to

choose methods which allow them best to perceive and know and, secondly, Stol et al. (2016) recommend selecting a primary data source as the basis for further data collection activities.

In the end, it was decided to select the Ambient Surfaces' custom interaction logging mechanism as the primary data source throughout the research. The motivation was fourfold:

- Firstly, interaction data allows one to shed light on a variety of usage patterns (e.g. content utilisation). Sensor data to track user activities is commonly utilised in ambient display research (Börner et al., 2013). Generally, logging is considered helpful in long-term enterprises (Alt et al., 2012).
- Secondly, phenomena such as the novelty effect and display blindness pose certain relevance (Koch et al., 2018). It was initially anticipated that a novelty effect would be present to some extent. Interaction data arguably allows the identification of uncommon patterns in the material.
- Thirdly, this method helped in keeping initial resources in check. For instance, some studies report that they extended their research due to the prevalence of an initial novelty effect (e.g. Gallacher et al., 2015; Hazlewood et al., 2011). Additionally, behavioural sciences show that it can take up to several months until a new behaviour takes hold (Prochaska and Diclemente, 1982). Data collection techniques such as observations would had arguably accounted for more time-intensive workloads (e.g. travel time), which is also mentioned by Corbin and Strauss (2015).
- Lastly, as it is crucial in ambient display research to collect data unobtrusively (Börner et al., 2013), the logging mechanism arguably allows for the collection of data without distracting potential users.

Initially, it was anticipated to sample data for at least a couple of months, primarily due to the novelty effect. Analyses were scheduled to be conducted weekly. Generally, data saturation is a crucial part of GT research (Corbin and Strauss, 2015). At its core, it develops—or saturates—the core categories of an emerging theory (Corbin and Strauss, 2015; Glaser, 2008). In Phase 1, data saturation referred to an incremental mitigation process that resulted from the decreasing impact of newly integrated interaction data on manifested latent patterns.

Findings of the Analysis

Below, findings obtained by continuously comparing data in Phase 1 are discussed. However, as this paper primarily concerns the methodological foundations of our research, this section exemplary presents some results stemming from this comparison process and indicates conclusions (i.e. hypotheses and emerging questions) drawn on its basis.

Analyses Stages

Overall, analyses in 2014 happened to be organised in four stages, while each of these stages concerned interaction data from periods of up to several weeks. This organisation is a result of particular events and observations. These four stages were:

- *Weeks 8 to 9 (Stage 1)*: As the Ambient Surface was operational on a Friday right before midday, it was decided to initiate analyses with the data from the first two weeks (i.e. six working days).
- *Weeks 10 to 11 (Stage 2)*: It was then decided to summarise the analysis activities from weeks 10 and 11, primarily as in Week 11 an update was deployed. Simultaneously, Week 11 marked the end of the reported novelty effect's overall existing time in some studies (e.g. Gallacher et al., 2015; Hazlewood et al., 2011).
- *Weeks 12 to 19 (Stage 3)*: Here, the analyses stood in light of the now arguably less prevalent effects of novelty towards the data. It was focused on a longer time period for two main reasons: firstly, while comparing the weeks in question, it was found that weeks 8 to 11 indicated the highest number of interactions; secondly, a five-week time period followed, in which the Ambient Surface was largely not operational due to software-related issues.
- *Weeks 20 to 52 (Stage 4)*: Of particular interest in this stage was whether the Ambient Surface could tackle the threats stemming from display blindness beyond a novelty effect and whether the previously obtained patterns endured over time. Consequently, it was decided to focus on analysing interaction data until the end of the year. The reason for this decision was threefold: firstly, to allow a profound comparison, it felt reasonable to collect a notable amount of data; secondly, as the Ambient Surface was lastly revised in Week 16 and the project partner was still occasionally reviewing its contents, it was hypothesised that more time had to pass to grasp on issues relating to display blindness; finally, the almost entirely non-operational state of the screen in weeks 20 to 24 posed a particular threat to display blindness as this issue could have resulted in disuse of the system afterwards.

Throughout all stages, memos in the form of what is depicted in Figure 3 were intensively being utilised. While cross-tabulations in Stage 1 provoked the idea that events such as arriving at work might play an important role relating to actual usage, Stage 2 revealed that changes to the system resulted in a measurable difference in the total number of interactions. With respect to display blindness, it was observable during Stage 3 that the Ambient Surface was still frequently being utilised. While the total number of interactions notably decreased, the system was seemingly incorporating some positive contributing factors. Otherwise, it was believed that the usage would have dropped more substantially or would have halted entirely. In Stage 4, it seemed that data became more representative compared to data stemming from preceding weeks. It was concluded that threats resulting from display blindness were likely to be prevalent to some extent, but it was also recog-

Daytimes	# of	# of (in %)
07:00	0	0.00%
07:30	0	0.00%
08:00	0	0.00%
08:30	0	0.00%
09:00	0	0.00%
09:30	0	0.00%
10:00	0	0.00%
10:30	0	0.00%
11:00	0	0.00%
11:30	0	0.00%
12:00	167	8.96%
12:30	161	8.64%
13:00	259	13.90%
13:30	302	16.21%
14:00	317	17.02%
14:30	27	1.45%
15:00	61	3.27%
15:30	5	0.27%
16:00	24	1.29%
16:30	208	11.16%
17:00	176	9.45%
17:30	36	1.93%
18:00	120	6.44%
18:30	0	0.00%
19:00	0	0.00%
19:30	0	0.00%
20:00	0	0.00%
Total:	**1,863**	**100.00%**

Daytimes	# of	# of (in %)
07:00	0	0.00%
07:30	47	1.22%
08:00	139	3.62%
08:30	155	4.03%
09:00	194	5.05%
09:30	83	2.16%
10:00	117	3.04%
10:30	293	7.62%
11:00	113	2.94%
11:30	146	3.80%
12:00	324	8.43%
12:30	184	4.79%
13:00	330	8.58%
13:30	213	5.54%
14:00	234	6.09%
14:30	159	4.14%
15:00	195	5.07%
15:30	86	2.24%
16:00	215	5.59%
16:30	128	3.33%
17:00	109	2.83%
17:30	125	3.25%
18:00	122	3.17%
18:30	97	2.52%
19:00	37	0.96%
19:30	0	0.00%
20:00	0	0.00%
Total:	**3,845**	**100.00%**

Daytimes	# of	# of (in %)
07:00	0	0.00%
07:30	47	0.82%
08:00	139	2.44%
08:30	155	2.72%
09:00	194	3.40%
09:30	83	1.45%
10:00	117	2.05%
10:30	293	5.13%
11:00	113	1.98%
11:30	146	2.56%
12:00	491	8.60%
12:30	345	6.04%
13:00	589	10.32%
13:30	515	9.02%
14:00	551	9.65%
14:30	186	3.26%
15:00	256	4.48%
15:30	91	1.59%
16:00	239	4.19%
16:30	336	5.89%
17:00	285	4.99%
17:30	161	2.82%
18:00	242	4.24%
18:30	97	1.70%
19:00	37	0.65%
19:30	0	0.00%
20:00	0	0.00%
Total:	**5,708**	**100.00%**

Figure 3. Three memos including the crude index *Utilisation of the Ambient Surface*, collating interaction data from Week 8 (left), Week 9 (centre), and weeks 8 and 9 in combination (right). While darker blue segments indicate stronger usage, white segments indicate the opposite.

nised that employees continued utilising the system. Therefore, the obtained latent patterns arguably became sufficiently saturated to seek additional material (e.g. observational data) to be included in the overall analysis and, hence, to strengthen theoretical sensitivity.

2014: A Retrospective

Before Phase 1 was concluded, two further investigations were retrospectively conducted: firstly, the all-year usage in terms of interactions per calendar week was elaborated and, secondly, data representativeness was analysed.

Comparing all of the interaction data from 2014, it was found that the mean number of interactions per week first fell below the all-year mean number of interactions in Week 18. We demonstrated this eleventh week into the field study elsewhere as the threshold which indicated that the initial novelty effect finally subsided (Koch et al., 2018). In investigating data representativeness, we intended to identify the individual interaction incidents that potentially affected conclusions drawn from analyses. In the end, it was not evident that the results during weeks 18 to 52 were substantially affected by any single incident. The Ambient Surface was used regularly and on most days at around 12:30 (see Figure 4).

Consequently, it was assumed that the findings showed 'naturalistic usage' (Siek et al., 2014) to a certain degree beyond both a novelty effect and threats relating to display blindness.

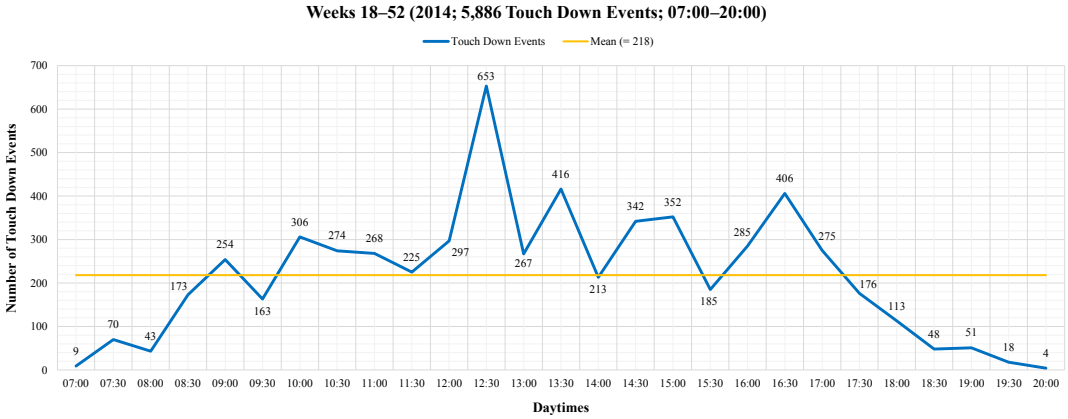


Figure 4. The total number of interactions (i.e. touch down events) with respect to corresponding daytimes for weeks 18 to 52 in 2014.

Conclusions

It is concluded with an illustration of hypotheses and questions which emerged throughout Phase 1. Both represented the basis for subsequent research phases and fostered theoretical sensitivity. Overall, the rather descriptive findings indicated that appropriation might be somewhat related to *informal* occasions (e.g. lunch breaks). In summary, the following hypotheses were posed:

- *The Ambient Surface is notably utilised in the early morning when people are arriving at work*
- *The Ambient Surface is mostly utilised when people are going to or are returning from lunch*
- *The Ambient Surface is notably utilised in the late afternoon when people are leaving work*
- *There are daytimes in which the use of the Ambient Surface notably decreases*
- *A novelty effect results in the usage of the Ambient Surface which distracts from latent patterns due to the magnitude of interactions and the time of their occurrence*
- *Changes to an existing system contribute towards or extend a prevalent novelty effect*
- *The Ambient Surface provides positive contributing factors beyond a prevalent novelty effect*

In addition to these hypotheses, Phase 1 revealed questions that largely surrounded the limitations of quantitative interaction data. These questions included:

- *What positive contributing factors is the Ambient Surface promoting?*
- *What are the reasons for the varying reductions in interactions throughout the day?*
- *How is the Ambient Surface passively utilised by staff members?*

- *How do the relationships of variables investigated in the analysis change when incrementally compared to additional interaction data?*
- *What can be learnt with respect to display blindness by utilising further data collection techniques?*
- *How do the conclusions regarding data saturation change when compared to additional interaction data from subsequent years?*

Discussion

The paper presents a practical systematisation of GT's two core concepts—constant comparison and theoretical sampling—in the early stages of our ongoing research. Generally, we do not argue that our approach is the modus operandi to conduct longitudinal ambient display in situ studies. For example, Siek et al. (2014) recommend a more sequential organisation of the research process, where analyses are carried out at the end of the field research. In fact, this post-deployment analysis approach is also chosen by some GT studies in the context of HCI and CSCW (Muller, 2014). Above, further utilised methodologies were also introduced (e.g. heuristic inquiries)—admittedly, other circumstances may seek a different methodological choice (e.g. limited time resources). In this respect, we seek to illustrate *a* way to methodologically guide in situ ambient display research.

Given that we are fundamentally interested in evaluating 'naturalistic usage' (Siek et al., 2014), we have committed to the challenge that research prototypes typically do not withstand daily use in authentic environments (Nunamaker Jr. et al., 2015; Siek et al., 2014). It is with the utmost certainty that the Ambient Surfaces would not be in operation as of 2019 had we not committed to this proactive engagement. Consequently, we faced several challenges throughout the entire study. As the Ambient Surfaces did and do change regularly (see Figure 5), so did and does the environment, including staff members, meeting schedules, holiday and illness seasons, as well as trainings—to name but a few. Following GT helped us to stay sensitive towards the data. For example, sometimes there were no interactions during an entire week. We immediately started to ask questions and postulate possible explanations. Here, the primary data source served as a sensitive indicator to rapidly start wondering about the data. We found this quantitative data source to be very helpful, especially when the research commenced but also later during the study (e.g. to isolate a novelty effect).

We find it difficult to convey the tacit knowledge regarding issues that are not directly presented in charts, diagrams, and statistical tests. However, as we are constantly comparing data and theorising about the implications, GT ensures that we explicitly track the progress while, for instance, asking questions and conducting analyses. Overall, by only considering such occasionally intertwined issues and by bringing them to the fore, we argue that field deployment reports reach their full potential and the reader is able to enjoy a text to its full extent.

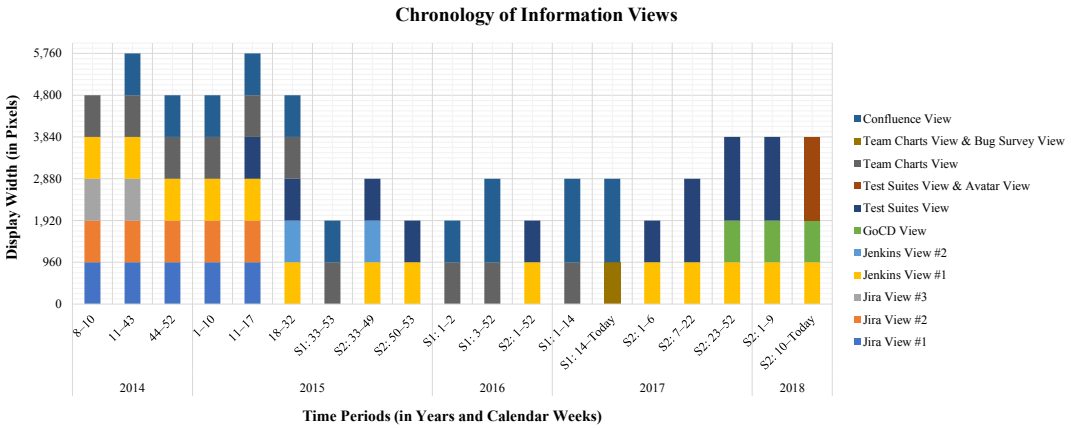


Figure 5. A bar chart diagram depicting the evolution of the Ambient Surfaces (i.e. 'S1' and 'S2' as of Week 33 in 2015) over time by relating added, removed, and resized (i.e. increases in width) information views to actual display widths (i.e. in pixels) and time periods.

Contributions

The following elaborates the different contributions this paper adds to existing knowledge, categorised in terms of field deployments, ambient displays, GT methodology, and the novelty effect.

Long-term Field Deployment Research

- Firstly, by shedding light on our ongoing long-term field deployment study, this paper responds to recent calls for more longitudinal in situ evaluations (Börner et al., 2013; Hazlewood et al., 2011; Preim et al., 2018).

Methodological Advances in Ambient Display In Situ Research

- Secondly, by envisioning the application of classic GT in ambient display field deployment research, the paper introduces a way to cope with the lack of methodological development in this domain (Hazlewood et al., 2011).

Grounded Theory Methodology

- Thirdly, it contributes to the general rare application of GT methodology in the domains of software engineering (Stray et al., 2016) as well as HCI and CSCW (Muller and Kogan, 2012).
- Fourthly, it demonstrates a practical systematisation of both core processes in GT. The literature remains silent as to how to proceed during the constant comparison process and fails to indicate what constitutes fertile comparison candidates to develop a theoretical model (Boeije, 2002). Stol et al. (2016) also note that the process of theoretical sampling remains unclear, particularly considering the implications of data magnitude in this process.

- Fifthly, it envisions the utilisation of quantitative data in GT research. There have been no recent attempts in this regard (Glaser, 2008).
- Sixthly, as the systematisation approach is used with both quantitative and qualitative data, this paper puts forward a means to use a mixed-methods approach in GT. Rarely are both data sources combined in GT research (Walsh, 2015).

Novelty Effect Research

- Seventhly and lastly, in addition to a discussion of selected examples of the novelty effect in a previous publication (Koch et al., 2018), this work illustrates how we coped with this effect on a methodological level.

Limitations

Overall, our study is conducted in one particular environment—other settings will likely reveal notable differences in terms of comparisons and the sampling strategy. Furthermore, as there is little practical guidance on conducting GT with quantitative data, it is possible that we misinterpreted certain parts that Glaser (2008) foresees for such research. Additionally, the issue of incorporating literature in GT research was disregarded (Giles et al., 2013), primarily due to the fact that we would have had to frame the paper more holistically. In a similar vein, the paper only briefly draws on the complexity of our research and does not convey any information on how the descriptive findings from Phase 1 transcended into conceptually representative categories. Also, social aspects such as how usage relates to practices (e.g. team meetings) are not further considered. Again, Phase 1 was intended as being the first stepping stone towards the goal of generating a substantive theory.

Conclusion and Recommendations for Future Work

In response to the lacking methodological advances in ambient display in situ research, this paper illustrates the application of classic GT in our ongoing longitudinal study. It seeks to put forward practical guidance for the researcher to go along with the methodology's two core concepts, namely constant comparison and theoretical sampling. To this end, it demonstrates a practical systematisation of both processes to compare and sample quantitative interaction data. The concept of crude indices was leveraged to unveil latent patterns and to indicate directions for future data collection as well as analysis activities on the basis of emergent hypotheses and questions.

In terms of future research, we propose the following directions. Firstly, a promising avenue would be to apply our approach to further quantitative analyses in related studies. Secondly, we encourage other researchers to conduct studies with GT as it has yet to gain momentum in HCI and CSCW research (Muller and

Kogan, 2012). Thirdly, irrespective of the selected research methodology, this paper puts forward the issue that in situ research is highly relevant (Börner et al., 2013; Hazlewood et al., 2011; Nunamaker Jr. et al., 2015; Preim et al., 2018; Siek et al., 2014). Consequently, any research with any chosen methodology conducted in this manner, would contribute valuable knowledge to the community. Fourthly and lastly, we concur with Bjørn and Boulus-Rødje (2015) and invite other researchers to rethink their research approaches when planning to conduct research in dynamic and heterogeneous environments. With the selection of a primary data source (e.g. to cope with the novelty effect), we hope to indicate some first stepping stones in this regard.

References

Alt, F., S. Schneegaß, A. Schmidt, J. Müller, and N. Memarovic (2012): 'How to Evaluate Public Displays'. In: *Proceedings of the 2012 International Symposium on Pervasive Displays*. New York, NY, USA, pp. 17:1–17:6, ACM.

Bannon, L. and C. O'Malley (1984): 'Problems in Evaluation of Human-Computer Interfaces: A Case Study'. In: B. Shackel (ed.): *Proceedings of IFIP INTERACT'84: Human-Computer Interaction*. Amsterdam, pp. 709–713, North-Holland.

Beck, K. (2000): *Extreme Programming Explained: Embrace Change*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.

Bjørn, P. and N. Boulus-Rødje (2015): 'Studying Technologies in Practice: "Bounding Practices" When Investigating Socially Embedded Technologies'. In: V. Wulf, K. Schmidt, and D. Randall (eds.): *Designing Socially Embedded Technologies in the Real-World*. London: Springer, pp. 341–356.

Boeije, H. (2002): 'A Purposeful Approach to the Constant Comparative Method in the Analysis of Qualitative Interviews'. *Quality and Quantity*, vol. 36, no. 4, pp. 391–409.

Börner, D., M. Kalz, and M. Specht (2013): 'Closer to You: Reviewing the Application, Design, and Evaluation of Ambient Displays'. *Int. J. Ambient Comput. Intell.*, vol. 5, no. 3, pp. 16–31.

Charmaz, K. (2006): *Constructing Grounded Theory*. Thousand Oaks, CA, USA: Sage.

Corbin, J. M. and A. L. Strauss (2008): *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Thousand Oaks, CA, USA: Sage, 3rd edition.

Corbin, J. M. and A. L. Strauss (2015): *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Thousand Oaks, CA, USA: Sage, 4th edition.

Crotty, M. J. (1998): *The Foundations of Social Research: Meaning and Perspective in the Research Process*. London: Sage Publications Ltd.

Dingsøyr, T., T. E. Fægri, and J. Itkonen (2014): 'What Is Large in Large-Scale? A Taxonomy of Scale for Agile Software Development'. In: A. Jedlitschka, P. Kuvaja, M. Kuhrmann, T. Männistö, J. Münch, and M. Raatikainen (eds.): *Product-Focused Software Process Improvement*. Cham, pp. 273–276, Springer International Publishing.

Gallacher, S., J. O'Connor, J. Bird, Y. Rogers, L. Capra, D. Harrison, and P. Marshall (2015): 'Mood Squeezer: Lightening Up the Workplace Through Playful and Lightweight Interactions'. In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. New York, NY, USA, pp. 891–902, ACM.

Giles, T., L. King, and S. de Lacey (2013): 'The Timing of the Literature Review in Grounded Theory Research An Open Mind Versus an Empty Head'. *Advances in Nursing Science*, vol. 36, no. 2, pp. E29–40.

Glaser, B. G. (1978): *Theoretical Sensitivity*. Mill Valley, CA, USA: Sociology Press.

Glaser, B. G. (1992): *Emergence versus forcing: Basics of grounded theory analysis*. Mill Valley, CA, USA: Sociology Press.

Glaser, B. G. (1998): *Doing Grounded Theory: Issues and Discussions*. Mill Valley, CA, USA: Sociology Press.

Glaser, B. G. (2006): *Doing formal grounded theory: A proposal*. Mill Valley, CA, USA: Sociology Press.

Glaser, B. G. (2008): *Doing Quantitative Grounded Theory*. Mill Valley, CA, USA: Sociology Press.

Glaser, B. G. and A. L. Strauss (1967): *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York, NY, USA: Aldine de Gruyter.

Hazlewood, W. R., E. Stolterman, and K. Connelly (2011): 'Issues in Evaluating Ambient Displays in the Wild: Two Case Studies'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 877–886, ACM.

Heath, H. and S. Cowley (2004): 'Developing a grounded theory approach: a comparison of Glaser and Strauss'. *International Journal of Nursing Studies*, vol. 41, no. 2, pp. 141–150.

Koch, M., K. von Luck, J. Schwarzer, and S. Draheim (2018): 'The Novelty Effect in Large Display Deployments – Experiences and Lessons-Learned for Evaluating Prototypes'. In: *Proceedings of 16th European Conference on Computer-Supported Cooperative Work - Exploratory Papers*, Vol. 2 of *Reports of the European Society for Socially Embedded Technologies*. European Society for Socially Embedded Technologies (EUSSET).

Matthews, T., T. Rattenbury, and S. Carter (2007): 'Defining, Designing, and Evaluating Peripheral Displays: An Analysis Using Activity Theory'. *Human–Computer Interaction*, vol. 22, no. 1-2, pp. 221–261.

Memarovic, N., I. Elhart, and E. Rubegni (2016): 'Developing a Networked Public Display System'. *IEEE Pervasive Computing*, vol. 15, no. 3, pp. 32–39.

Messeter, J. and D. Molenaar (2012): 'Evaluating Ambient Displays in the Wild: Highlighting Social Aspects of Use in Public Settings'. In: *Proceedings of the Designing Interactive Systems Conference*. New York, NY, USA, pp. 478–481, ACM.

Morse, J. M., P. Noerager Stern, J. Corbin, B. Bowers, A. E. Clarke, and K. Charmaz (2009): *Developing Grounded Theory: The Second Generation*. London: Routledge.

Muller, M. (2014): 'Curiosity, Creativity, and Surprise as Analytic Tools: Grounded Theory Method'. In: J. S. Olson and W. A. Kellogg (eds.): *Ways of Knowing in HCI*. New York, NY: Springer, pp. 25–48.

Muller, M. J. and S. Kogan (2012): 'Grounded Theory Method in Human-Computer Interaction and Computer-Supported Cooperative Work'. In: J. A. Jacko (ed.): *The Human Computer Interaction Handbook*. Boca Raton, FL, USA: CRC Press, 3rd edition, pp. 1003–1024.

Nunamaker Jr., J. F., R. O. Briggs, D. C. Derrick, and G. Schwabe (2015): 'The Last Research Mile: Achieving Both Rigor and Relevance in Information Systems Research'. *Journal of Management Information Systems*, vol. 32, no. 3, pp. 10–47.

Ojala, T., V. Kostakos, H. Kukka, T. Heikkinen, T. Linden, M. Jurmu, S. Hosio, F. Kruger, and D. Zanni (2012): 'Multipurpose Interactive Public Displays in the Wild: Three Years Later'. *Computer*, vol. 45, no. 5, pp. 42–49.

Peltonen, P., E. Kurvinen, A. Salovaara, G. Jacucci, T. Ilmonen, J. Evans, A. Oulasvirta, and P. Saarikko (2008): 'It's Mine, Don'T Touch!: Interactions at a Large Multi-touch Display in a City Centre'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 1285–1294, ACM.

Preim, B., T. Ropinski, and P. Isenberg (2018): 'A Critical Analysis of the Evaluation Practice in Medical Visualization'. In: A. Puig Puig, T. Schultz, A. Vilanova, I. Hotz, B. Kozlikova, and P.-P. Vázquez (eds.): *Eurographics Workshop on Visual Computing for Biology and Medicine*. Granada, Spain, pp. 45–56, The Eurographics Association.

Prochaska, J. and C. Diclemente (1982): 'Transtheoretical therapy: Toward a more integrative model of change'. *Psychotherapy: Theory, Research & Practice*, vol. 19, no. 3, pp. 276–288.

Rogers, Y., W. R. Hazlewood, P. Marshall, N. Dalton, and S. Hertrich (2010): 'Ambient Influence: Can Twinkly Lights Lure and Abstract Representations Trigger Behavioral Change?'. In: *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*. New York, NY, USA, pp. 261–270, ACM.

Schwaber, K. and M. Beedle (2001): *Agile Software Development with Scrum*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1st edition.

Schwarzer, J., S. Draheim, K. von Luck, Q. Wang, P. Casaseca, and C. Grecos (2016): 'Ambient Surfaces: Interactive Displays in the Informative Workspace of Co-located Scrum Teams'. In: *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. New York, NY, USA, pp. 69:1–69:4, ACM.

Shelton, B. and K. Nesbitt (2017): 'Evaluating WaveWatch: An Ambient Display of Web Traffic'. In: *Proceedings of the Australasian Computer Science Week Multiconference*. New York, NY, USA, pp. 9:1–9:9, ACM.

Siek, K. A., G. R. Hayes, M. W. Newman, and J. C. Tang (2014): 'Field Deployments: Knowing from Using in Context'. In: J. S. Olson and W. A. Kellogg (eds.): *Ways of Knowing in HCI*. New York, NY: Springer, pp. 119–142.

Stern, P. N. (1995): 'Eroding grounded theory'. In: J. M. Morse (ed.): *Critical issues in qualitative research methods*. Thousand Oaks, CA, USA: Sage, pp. 212–223.

Stol, K. J., P. Ralph, and B. Fitzgerald (2016): 'Grounded Theory in Software Engineering Research: A Critical Review and Guidelines'. In: *2016 IEEE/ACM 38th International Conference on Software Engineering (ICSE)*. New York, NY, USA, pp. 120–131, ACM.

Strauss, A. L. (1987): *Qualitative Analysis for Social Scientists*. New York, NY, USA: Cambridge University Press.

Strauss, A. L. and J. M. Corbin (1990): *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Newbury Park, CA, USA: Sage.

Strauss, A. L. and J. M. Corbin (1998): *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Newbury Park, CA, USA: Sage, 2nd edition.

Stray, V., D. I. K. Sjøberg, and T. Dybå (2016): 'The daily stand-up meeting: A grounded theory study'. *The Journal of Systems and Software*, vol. 114, no. April 2016, pp. 101–124.

Walsh, I. (2015): 'Using quantitative data in mixed-design grounded theory studies: an enhanced path to formal grounded theory in information systems'. *European Journal of Information Systems*, vol. 24, no. 5, pp. 531–557.

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, pp. 111–150.

O'Keeffe, M.; Hogan, T.; Delaney, K. (2019): Revisiting and Rethinking the Structural Elements of a Community of Practice. In: Proceedings of the 17th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing an the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep11

Revisiting and Rethinking the Structural Elements of Communities of Practice

Michelle O'Keeffe, Trevor Hogan & Kieran Delaney
Human Data Interaction Group, Cork Institute of Technology, Cork, Ireland
michelle.m.okeeffe@mycit.ie, {trevor.hogan; kieran.delaney}@cit.ie

Abstract. Communities of Practice have existed for as long as people have been learning and sharing their experiences. However, it was not until the early 1990's before the study of these communities gained attention from the research community. Since then, these communities have been studied in many research domains, yet, the core structural elements, which are critical to these communities remain constant - *Domain, Community* and *Practice*. In this paper we re-examine the structural elements of Communities of Practice and argue for the extension of these to include aspects on *Participation*, *Learning* and *Knowledge*. We also take a first step in validating these new structural elements by presenting a study that explores how they appear in a known Communities of Practice (the CoderDojo movement). Our research informs the future study of COP from both a theoretical and organizational perspective.

Introduction

Communities of Practice (COP) have existed for as long as people have been learning and sharing their experiences through storytelling (Lave and Wenger, 1991). However the study of COP, and in particular their impact on how groups of people work and learn together only gained attention from research communities in the early-1990s. Today, COP are widely studied in many domains, such as Healthcare (cf. (Falkman, et al., 2008)), Software Engineering (cf. (Ranmuthugala, et al., 2011)) and Business and Management (cf. (Wenger and Snyder, 2000)) to name but a few. Over this period, much research has validated and confirmed Lave and Wenger's claim (1991) that the three main structural elements of a COP are:

Domain, Community and *Practice* (Corso and Giacobbe, 2005; Sánchez-Cardona, et al., 2012; Snyder, et al., 2003; Wenger, 1998; Wenger, 1998b; Wenger, 2006; Wenger, et al., 2002; Wenger and Snyder, 2000). Research also supports the assertion that these three elements are vital to the formulization and development of a COP (Wenger, 1998) and the combination of these three elements is what differentiates a COP from a normal community (Eckert, 2006).

The aim of our research is to revisit the structural elements of COP (Domain, Community and Practice) and in doing so also pose the questions: *What defines a Community of Practice? What are the key characteristics of a Community of Practice?* And *how do the key characteristics materialize in an existing Community of Practice?* Our objective is to explore previous research on COP to seek whether other structural elements have been discussed and could possibly be given as much prominence as the three that Lave and Wenger (1991) first established. In doing so, the approach we took involved a two-step process, whereby we first conducted an extensive review of literature related to COP, the results of which lead us to propose amending the three current structural elements (*Domain, Community* and *Practice*) with *Participation*, *Learning* and *Knowledge* to form a more holistic understanding of the structural elements of current COP. As a first step in validating this claim we conducted a follow-on study that involved embedding ourselves in a commonly known COP (the CoderDojo movement) whereby we conducted a series of semi-structured interviews with CoderDojo mentors to identify how and where these elements (*Participation, Learning* and *Knowledge*) appear in the day-to-day running of this community. The main contribution of this paper is twofold, first we present a review of research on COP, conducted in many domains over the last twenty-five years. Secondly, based on our review we propose an extension of the current structural elements of COP and take a first step in validating these new elements.

Communities of Practice

Communities of Practice (COP), as a term and research area, was first introduced by Lave and Wenger (1991) when they discussed it in regards to apprenticeship as a learning model. Wenger then extended this concept, and applied it to other domains, such as, for instance: organisations (Wenger, 1998b). As part of this work Wenger defined COP as a "group of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly" (1991, pp 72-73). Many others have developed different definitions of COP in relation to their use in various types of organisations and sectors (Brown and Duguid, 1991; Hildreth & Kimble, 2002; Fisher, 2001; Hoadley, 2012; Lindkvist, 2005; Gherardi, 2006).

While there are many different definitions of COP in the literature, there is more consensus on the key structural elements that help to form and develop a COP. Lave

and Wenger (1991) first discussed the three structural elements of domain, community and practice in their early work and they were further elaborated by Wenger (1998). They first introduced them to describe the way in which a COP can evolve naturally because of the members' common interest and commitment to a particular domain or area. The early work of Lave and Wenger, which identified these elements, has subsequently been developed further throughout the years. If we look at each element individually, *Domain* refers to the shared domain of interest the community members are engaged in (Corso and Giacobbe, 2005; Probst and Stefano, 2008; Ranmuthugala, et al., 2011; Wiggberg and Daniels, 2011) and commitment to this domain is implied by membership and a shared competence that distinguishes members from other people (Lave and Wenger, 1991; Wenger, 1998, 1998b, 2006). *Community* refers to the joint activities the community members engage in (Lave and Wenger, 1991; Wenger, 1998, 1998b, 2006), i.e. they share information, help each other, engage in group discussions, build relationships etc. Finally, *Practice* refers to the shared repertoire of resources the community members build up, (Lave and Wenger, 1991; Wenger, 1998, 1998b, 2006), i.e. stories, tools, experiences, ways of dealing with recurring problems, etc. These elements have been validated by many researchers in the current COP literature (see for instance, Ranmuthugala, et al., 2011; Nistor and Fischer, 2012; Sánchez-Cardona, et al., 2012; Corso and Giacobbe, 2005; Snyder and Briggs, 2003; Lindkvist, 2005; Wiggberg and Daniels, 2011).

Our research leverages on these works, as well as those that sought to explore structural elements beyond those first established by Lave and Wenger (1991). Couros and Kesten (2003) explored COP by conducting a considerable literature review to elicit the different definitions of a COP that exist, along with what they call the "general characteristics" of a COP (Lave and Wenger, 1991, pp. 10). Under this term they discuss characteristics such as social learning theory, participation, knowledge sharing, and more. However, they don't go so far as to describe or define these as key characteristics or structural elements of a COP. Also, the literature they used to identify these characteristics may be considered as being quite narrow as they mainly relied on the work of Wenger (1998; 1998b). Since then, the body of literature on COP has greatly expanded.

As part of a review of literature on COP, Roberts (2006) explored the role of a COP for interpersonal knowledge transfer. In this she summarised the key characteristics of a COP compiled by Wenger (1998b), however, she overlooked the three aspects we focus on in this paper: *Participation*, *Learning* and *Knowledge*. The characteristics she did focus on can easily be identified with *Domain, Community* and *Practice*. Although Roberts does mention participation with regards to negotiating meaning (Roberts, 2006, pp. 4) and forming an identity within the COP (pp. 8) she does not explain how one participates in a COP and fails to highlight the importance of participation as a key characteristic.

In this section we provided a brief overview of research that has sought to define and elaborate on the study of COP. In the next section we continue our review of related literature but we focus on work that has explored the structural elements of COP with the aim of identifying key elements that may have been somewhat overlooked in previous work.

Reviewing and Extending the Structural Elements of COP

This phase of the study followed on from the initial review of literature on COP. We decided to conduct a Meta Review of the literature in order to extract the most prominent structural elements of a COP. In doing so, our aim was to explore whether or not there are structural elements discussed in the literature that may have emerged since Lave and Wenger (1991) first established the three known elements of Domain, Community and Practice - and if so, why they have not been considered as structural elements before. This phase involved reviewing literature that investigated different aspects of COP. We extracted the definitions and characteristics from the literature and stored them in an online digital archive, keeping a detailed record of each. Next, we categorised these based on keywords and phrases related to the structural elements of Domain, Community and Practice. During this process, we found many characteristics could not easily be categorised under these elements, including the way in which a member of a COP moves from the periphery to the centre through participation, the way in which knowledge is transferred from one member to another, and the type of learning that takes place in a COP. Therefore we went through a process of coding, categorising and describing the characteristics that did not relate to Domain, Community and Practice and we ended up with a set of three new structural elements: Participation, Learning and Knowledge. In Table 1 we present a list of characteristics under these elements that were prominent in the literature. We categorised all of the similar characteristics in the literature together based on similarities in key phrases and meaning. In the following we discuss these structural elements in more detail.

Participation

According to the literature, participation in a COP refers to the accumulation of expertise, the development of expertise and the stimulation of the social construction of knowledge (Nistor and Fischer, 2012). Members in a COP go through a process of legitimate peripheral participation (Couros and Kesten, 2003; Lave and Wenger, 1991; Wenger, 1998b). People initially join the community and learn on the periphery, and as they become more competent, and the activities of the community become more relevant, they move closer to the core of the

community (Hoadley, 2012; Lave and Wenger, 1991; Wenger, 1998b). Through this process of legitimate peripheral participation the COP members gain expertise with the COP and they construct knowledge. One's identity in a COP can be described in terms of various levels of expertise. Members start as a novice on the periphery moving towards an expert at the centre of the community. This expert status comes as a result of participation (Nistor and Fischer, 2012).

Table 1 – COP Characteristics - extracted from existing literature under the structural elements Participation, Learning and Knowledge.

Participation	**Learning**	**Knowledge**
Legitimate Peripheral Participation *(Eckert, 2006; Lindkvist, 2005; Wenger, 2006)*	Informal Learning *(Boud, Middleton, 2003; Wenger, 1998b)*	Tacit and Explicit Knowledge *(Avasti, et al., 2015; Bradshaw et al., 2004; Sauve, 2007; Wenger, 2006)*
Members join and learn on the periphery of the community before moving towards the center *(Hoadley, Kilner, 2005; Lindkvist, 2005; Wenger, 2006)*	Learning is an active process *(Lindkvist, 2005)*	Members develop advanced and reproducible knowledge in the community domain *(Probst, Borzillo, 2008)*
Members gain experience with the practice of the community *(Probst, Borzillo, 2008)*	Social Participation *(Lindkvist, 2005; Wenger, 1998; Wenger, Snyder, 2000)*	Knowledge is categorized by narratives, collaboration and constructivism *(Blankenship, Ruona, 2007)*
Member construct knowledge through participation *(Probst, Borzillo, 2008)*	Participation fosters learning in a COP *(Lindkvist, 2005)*	Knowledge is transferred through informal methods *(Sheridan, Goggin, O'Sullivan, 2016)*
Members build up and develop their expertise within the COP *(Avasti, et al., 2015, Eckert, 2006; Probst, Borzillo, 2008; Wenger, 2006)*	The person is actively involved in their learning process *(Lindkvist, 2005; Wenger, 1998; Wenger, Snyder, 2000)*	Knowledge in a COP is created and disseminated through the C4P model *(Jakovljevic, et al., 2013)*
Participation leads to various levels of expertise & expert status in a COP *(Probst, Borzillo, 2008; Wenger, 2006)*	Situated Learning Theory *(Clancey, 1995; Lindkvist, 2005; Wenger, 1998)*	Narratives or stories are used to identify problems and represent existing knowledge *(Jakovljevic, et al., 2013)*
Members start as a novice and through participation become an expert *(Probst, Borzillo, 2008; Wenger, 2006)*	Learning takes place in the same situation it is applied *(Wenger, 1998; Wenger, Snyder, 2000)*	Collaboration involves members engaging in and sharing common practice *(Jakovljevic, et al., 2013)*
Expertise can determine the level of participation a member engages in *(Probst, Borzillo, 2008)*	Members learn through content, context, community and participation *(Clancey, 1995; Lindkvist, 2005; Wenger, 1998)*	Constructivism allows members to develop an understanding of the practice and how to solve problems *(Jakovljevic, et al., 2013)*
Different levels of participation – (i) Core group,	Formal learning can sometimes take place	Collaboration and interaction increases and improves

(ii) those who engage in discussions & activities and (iii) periphery group *(Wiggberg, Daniels, 2011)*	through training sessions, workshops and courses *(Sauve, 2007)*	knowledge in a COP *(Jakovljevic, et al., 2013)*

Learning

The process of legitimate peripheral participation is a social process and fosters learning in a COP (Lave and Wenger, 1991). Lave and Wenger describe it as follows: 'learning occurs if the person is actively involved in the learning processes'. This process of learning is grounded in the situated learning theory (Clancey, 1995; Lave and Wenger, 1991; Stein 1998) and Lave and Wenger relate this to the learning within a COP. Members of the COP acquire professional skills and the process of legitimate peripheral participation leads to full memberships in a COP (Lave and Wenger, 1991). Situated learning is the relationship between learning and the social context in which it occurs, i.e. learning takes place in the same situation in which it is applied (Stein 1998; Wenger, et al., 2002). The major elements in situated learning are: content (facts and processes of a task); context (situations, values, environment, and cues); community (the group where the learner will create and negotiate) and participation (where a learner works together with others in order to solve the problem) (Clancey, 1995; Lave and Wenger, 1991; Stein 1998). This form of learning is a social process and comes as a result of social participation, i.e. the individual is an active member of the COP and in the construction of their identity within the community (Stein 1998; Wenger, et al., 2002). According to Wenger (1998), social participation leads to the informal learning which takes place in a COP. This form of learning takes place within the informal meetings, conversations and relationships the community members engage in (Boud and Middleton, 2003).

Knowledge

Communities of practice are very important sources of knowledge (Avasthi, et al., 2005), and it has been argued that they have become the most natural way in which tacit knowledge is transferred within organisations (Bradshaw, et al., 2004). The types of knowledge created and disseminated in a COP can be described as tacit and explicit knowledge (Avasthi, et al., 2005; Sánchez-Cardona, et al., 2012; Wenger, 1998b). COP can be used to facilitate the informal knowledge transfer that drives leadership development, productivity, and innovation because the amount of work driven by tacit knowledge continues to rise (Sauve, 2007). The creation of knowledge in a COP can be categorised by the following three elements – Narratives, Collaborations and Social Constructivism (Blankenship and Ruona, 2007; Brown and Duguid, 1991; Couros and Kesten, 2003; Hoadley, 2012). Narratives are used for identifying problems and representing existing knowledge.

Collaboration refers to participants engaging in and sharing common practice. Social Constructivism describes how participants develop a common understanding of their practice and how to solve problems. The C4P Framework (Content, Conversations, Connections, Context and Purpose) established by Hoadley and Kilner (2005) describes how knowledge is created and disseminated in a COP.

We argue that our proposed structural elements - *Participation*, *Learning* and *Knowledge* - alongside *Domain*, *Community* and *Practice,* provide us with a more comprehensive way of describing and understanding how COP form and develop. These extra elements may not have been considered as key elements before, as the existing literature has arguably sought to confirm Lave and Wenger's (1991) postulation of three key elements. Lave and Wenger (1991) identified Domain, Community and Practice as structural elements of a COP and these elements have a number of links between them. For example, the domain is what brings the community together and if the community doesn't interact on a regular basis, they cannot develop their practice. Lave and Wenger (1991) discussed characteristics of participation, learning and knowledge within the context of a COP and highlighted their importance but did not consider them as structural elements. This may be because they chose to discuss them within the context of the Domain, Community and Practice and within the context of how a COP operates.

In the following, we present a study aimed at validating the extension of the structural elements. We focused on how and where aspects of Participation, Learning and Knowledge appear in a recognised COP - the CoderDojo Movement.

The CoderDojo Study

Before we describe the study, we will describe the CoderDojo movement, its structure, as well as the rationale for us selecting it as an exemplar COP to study. The CoderDojo movement is a global network of public, volunteer-led, community-based programming clubs for young people aged 7-17 years old. Through participation in these clubs, young people learn how to code, build a website, create an app or a game and explore technology in an informal, creative and social environment. This community is relatively new, starting first in Cork, Ireland in July 2011. It soon became a global phenomenon, with over 10,000 registered coding clubs in 125 countries worldwide[1] and the growth of several new Dojos each week.

[1] https://www.codeclubworld.org/

The CoderDojo Structure

The CoderDojo is a global community of coding experts, parents and children who all come together for one purpose - to teach and learn how to code in an informal learning environment. The core group of a CoderDojo includes the mentor - typically the person who set up the CoderDojo (i.e. CoderDojo Champion) - and the children who attend the CoderDojo. Each child is initially on the periphery of the community where they learn how to code, but as they become more proficient at coding and familiar with the activities of the community, they move towards the core of the community by helping newcomers and becoming junior mentors. The parents are generally on the periphery (unless they set up the dojo), although they also have an opportunity to learn how to code with their child. Each individual Coderdojo can be viewed as a COP – however, each dojo has little or no interaction with each other. Instead they could be viewed as a network of COP which fall under the umbrella of the Coderdojo Foundation.

We selected the CoderDojo movement as an exemplar COP to study, not only because it is a commonly recognised COP, but also because we have access to over 20 local CoderDojos, which helped us to gather a number of different opinions and insights on the day-to-day running of the community. Access to the CoderDojos allowed us to understand how the CoderDojo mentors and participants communicate and collaborate with each other both inside and outside of their weekly clubs. Moreover, we anticipated that interviewing members of the CoderDojo would help us to understand what kind of learning outcomes are expected or anticipated from participation in the CoderDojo.

The Study

The aim of this part of the study was to investigate whether or not the new structural elements established as part of phase one, i.e. *Participation, Learning* and *Knowledge,* appear as key elements in the day-to-day running of a CoderDojo. With this in mind we collected in-depth insights into the everyday running of the Coderdojo from those at the core of the community. We conducted ten semi-structured interviews with ten CoderDojo mentors (P1-10) from eight different Coderdojo groups (C1-8) (See Table 2.) The interviews were conducted with 6 female and 4 male mentors and the age group varied between 17 and 60. Two participants are CoderDojo Champions, i.e. they are responsible for setting up the dojo but do not mentor the participants. Five participants are CoderDojo Champions and Mentors, i.e. they were responsible for setting up the dojo and they also mentor the students. Another participant is a CoderDojo Mentor only – i.e. she attends the dojo each week and mentors the students with her background in coding. Finally, two participants are CoderDojo Champions and parents to some of the participants, i.e. they set up the dojo because they thought it would be a good activity for their children. These participants are also involved in the logistics of

the CoderDojo, i.e. organising the venue, trips, guest speakers, etc. Each interview lasted between 25 and 40 minutes. The questions asked were open-ended and focused around aspects that relate to Participation, Learning and Knowledge to determine if they were prominent characteristics in the day-to-day running of their CoderDojo. The questions asked included: How did you get involved in the CoderDojo and why? What have you learnt during your time in this CoderDojo? Tell me about the activities of this CoderDojo – can you describe a typical CoderDojo class? All interviews were video and audio recorded, and the same researcher who conducted the interviews transcribed them (lead author).

Table 2 – Interview Participant Information

Interview Participant/ Coderdojo	**Coderdojo Venue**	**Mentors & Participants in the Coderdojo**	**Champion**	**Mentor**	**Parent**
P1/C1 (male)	Community Library	20 participants; 2-4 mentors	✓	✓	×
P2/C2 (male)	Large Multinational Company	30 participants, 2-4 mentors	✓	✓	×
P3/C3 (male)	Large Multinational Company	30 participants, 2-4 mentors	✓	✓	×
P4/C4 (male)	Large Multinational Company	30 participants, 2-4 mentors	✓	✓	×
P5/C5 (female)	Community Centre	20 participants; 2-4 mentors	✓	×	✓
P6/C5 (female)	Community Centre	20 participants; 2-4 mentors	✓	×	✓
P7/C1 (female)	Community Library	20 participants; 2-4 mentors	×	✓	×
P8/C6 (female)	Community Hall	20 participants; 2-4 mentors	✓	×	×
P9/C7 (female)	Large Multinational Company	30 participants, 2-4 mentors	✓	✓	×
P10/C8 (female)	Third Level College	60 participants, 2-4 participants	✓	×	×

Data Analysis

As part of our analysis we coded the interview transcripts against the three new structural elements of a COP – Participation, Learning and Knowledge. We chose a directed content analysis approach (Bengtsson, 2016; Elo & Kyngäs, 2008;

Graneheim & Lundman, 2004, Helgevold & Moen, 2015) and close-coding approach as we sought to focus the analysis on aspects of the three new structural elements. The interview transcripts were read through several times to obtain a clear understanding of the data before the transcripts were divided into high level categories of Participation, Learning and Knowledge. Each of these transcript sections was then coded, which resulted in 45 unique codes (13 Participation, 20 Learning and 12 Knowledge). These codes were then sorted semantically whereby two themes emerged for each structural element (see Table 3.). These themes are discussed in the following section.

Table 3 – Data Analysis – Structural Elements, Themes and Codes.

Structural Element	**Theme (Unique Codes)**	**Exemplar Code**
Participation	• Legitimate Peripheral Participation (8) • Active Participation accumulates Expertise (5)	• Children helping their fellow participants • Collaborative participation
Learning	• Participation fosters learning (11) • Participants formulate their own learning (9)	• Children learning from the mentors • Children can personalise projects
Knowledge	• Participation leads to increased knowledge (6) • Knowledge is disseminated through collaboration & communication (6)	• Children increasing their knowledge of coding through participation • Face-to-face communicate & collaborate

Findings

Through our analysis of the interview transcripts, we identified numerous occasions where our participants referred to aspects of Participation, Learning and Knowledge when talking about their everyday experience of the CoderDojo. In the following we outline where these structural elements are evident in the day-to-day activities of the CoderDojo.

Aspects of Participation

Legitimate Peripheral Participation

Evidence from the interviews indicates that the participants of the CoderDojo (parents, children and mentors) go through a process of legitimate peripheral participation (Lave & Wenger, 1991). The parents bring their children to the

individual CoderDojo clubs, sitting on the periphery of the community but they often learn with their child as P2/C2 explains, *"Parents have an opportunity to have a good learning experience as well. They can learn the challenge and watch their child learn too. They engage in a joint exercise."* Parents can also move through the layers of the community as they become mentors to the new members. They gain experience with the community practice through learning and teaching coding. P6/C5 highlights this point, *"Parents end up being mentors because they pick up things along the way."* The children also join the community at the periphery, attending their first class and learning the basics of coding. They start with simple coding, moving towards more advanced coding as their coding expertise increases. In this way, they move through the layers of the community. As they become more proficient at coding, they move towards the centre of the community and act as a mentor to the new CoderDojo members: *"Kids do the mentoring of the blue belt. To get the blue belt, the idea is that they do three consecutive stints of teaching a class or group of students. The kids do an entire run of a class"* (P5/C4); *"We have one junior mentor and she is in the teenage room. She started here as a participant. Now she has started her own dojo in her own school"* (P9/C7). Mentors in the CoderDojo are immediately part of the central activities of the community. Some join at the centre from the beginning (e.g. those with a prior knowledge of coding) while others, through legitimate peripheral participation, move towards the centre and become mentors over time: *"I was going to CoderDojo as a participant for about six years since it started. Bill Lau who set up the original CoderDojo spoke to me and had been very encouraging about setting up my own one…I set up this CoderDojo when I was in Transition Year. My role is Mentor."*(P1/C1)

Active Participation accumulates Expertise

The children accumulate expertise of coding and problem solving as they actively participate in the classes - listening to their mentor, collaborating with fellow participants, asking questions, etc. This is how they construct their knowledge within the community: *"The motto is if you have a problem, ask three other kids first before you ask the mentor and this encourages group work and collaboration"* (P10/C8) and *"We (the mentors) tackle lots of questions asked by the kids. We tackle things line by line…..We do gentle quizzes to make sure the kids understand what is happening because the point is to learn how to code."* (P7/C1) It can be said in the CoderDojo that one's identity can be described as various levels of expertise, and that this expertise comes as a result of participation in the community. The parents and children typically join the community as novices, and some parents remain a novice on the periphery. Meanwhile, children move from being a novice to an expert at the centre of the community as they increase their knowledge of coding. Some children and parents become mentors to the younger children, and mentors are seen as experts in the community. The organiser of the

CoderDojo may not always be a mentor, although they are still central to the activities of the dojo. They are responsible for recruiting mentors and participants and organising events or activities for the dojo if necessary: *"I started bringing my son to this CoderDojo when it started. After about two weeks, I became secretary on the committee. We make sure the kids work on their projects…We try to organise day trips and guest speakers to come in when mentors are unable to attend."* (P3/C3).

Aspects of Learning

Participation fosters Learning

The transcripts reveal that legitimate peripheral participation also fosters the learning that occurs in a CoderDojo community. The participants learn by being actively involved in the activities of the dojo and by actively moving through their community. The participants are engaged in an informal learning environment, which is somewhat different from what they are familiar with at school, *"The dojo has a relaxed learning environment. If it's too like school, they won't want to come back again….They (the kids) might have their own thing that they would like to add, something they've seen on YouTube. The kids will do their own thing, changing variables in the code, etc."* (P7/C1) The CoderDojo facilitates a learning environment where the children can have fun and enjoy themselves while learning to code: *"We facilitate a learning environment. I want children to have fun and enjoy themselves. I want them to learn how to solve problems, something they can't learn in schools."* (P2/C2) The activities of the CoderDojo lead to informal learning activities that in turn facilitate collaborations and an exchange of knowledge between its participants. Children work on their own as well as in groups: *"Generally kids work on their own projects. Although they are working on their own games, they still work together. A lot of the kids teach themselves. They generally come up with their own ideas"* (P8/C6). They are also encouraged to help each other: *"The mentors go around and help the children and we also encourage the children to help each other….I love when the children are inquisitive and cheeky. I like a lively session! I love when kids ask questions"* (P9/C7).

Participants formulate their own Learning

The children are often the determinants of their own learning, and can influence the topics that are covered in the classroom, *"…we were doing tables but the kids didn't want to do it so we skipped past it. The kids can come up with ideas. For example, one kid wanted to add a video to his website so we dedicated a class to that."*(P1/C1) The mentors and the children acquire new skills as they participate in the CoderDojo, both technical and soft skills: *"The mentors learn communication skills, presenting skills and organisation skills"* (P5/C4). Meanwhile, along with coding, the children learn *"social, soft skills,*

communication, teamwork and people skills" (P9/C7). Some children will chose to work on their own projects, *"The children downstairs work on their own projects. They've been invited to take part in classes but they like working on their own stuff. They help each other*" (P4/C4). Mentors learn what's involved in running a dojo and what it means to be a mentor: *"I had to grasp the difference between teaching students at University College Cork (UCC) and kids who may be as young as 10. The students in UCC would pick things up a lot quicker…Adapting to the young age of the children was a challenge"* (P4/C4). Younger and older children learn different coding concepts: "*We introduce them (young children) to Scratch. That is used to introduce the children to coding, but the older kids think this is very juvenile. We try to give them the opportunity to work with Internet devices such as Raspberry Pi and others. Some kids learn Python and Java…"* (P2/C2).

Aspects of Knowledge

Participation leads to increased Knowledge

The content and approach to learning in a CoderDojo varies from one dojo to another. Some CoderDojo use PowerPoint slides and the CoderDojo book as sources of content for each class: *"It is my responsibility to prepare the material, know the code beforehand, prepare slides, PowerPoints, PDFs, teach the kids, explain the code to them and explain what it is used for"* (P1/C1). While other dojos will introduce the children to different software programmes and allow the children to work on their own projects: *"We shifted towards an overview of the environment (e.g. Scratch) for fifteen minutes and then a challenge from the first opportunity. We offer new challenges, rewards, demos but a lot of the projects in the class can be similar."* (P2/C2) The mentors and children engage in face-to-face conversations on a regular basis as they meet for classes each week.

Knowledge is disseminated through collaboration & communication

The mentors help the children with any problems they are facing, and in doing so disseminate knowledge to the community members. Two mentors we interviewed also use an online platform where mentors can discuss what they are delivering each week, and the children can ask questions of the mentors if they are working on a project outside of the class: *"We use a Slack Channel which is a team chat channel. There are different channels for different classes. On the mentors channel we discuss what we are going to cover each week. Students can ask questions about something they are stuck on."* (P5/C4), and *"We also have the Discord Channel where the kids can ask questions about anything to do with CoderDojo coding or separate coding."* (P7/C1). The children form friendships with each other and they also form a relationship with their mentors as they meet on a weekly basis and this forms trust between the community members, *"I wanted continuity. That way I knew what the child was working on each week and I could attempt to help the child*

the following weeks. We needed to build a relationship with the children" (P2/C2) and *"The friendships they make is nice to see as well. They help each other with the code, if one is finished they go around and help each other"* (P7/C1).

CONCLUSION & FUTURE WORK

It is evident in the literature that COP are widely studied across different sectors, yet only three structural elements of COP – *Domain*, *Community* and *Practice* – are used when identifying the key characteristics of a COP. We argue that these structural elements overlook how and why people participate in a COP. They also overlook aspects of learning that takes place within a COP and they don't represent how knowledge is created, managed and disseminated in a COP. During our review of COP literature we found that these characteristics were not listed under the main identifying features of a COP, yet our study of the CoderDojo movement seems to indicate that they are crucial to the formation, development and day-to-day running of this COP.

In summarising our work, we firstly presented a comprehensive review of research on COP, conducted in many domains over the last twenty-five years. Secondly, based on our review we proposed an extension of the current structural elements of COP and we took a first step in validating these new elements by investigating how these new structural elements appear in the day-to-day running of the CoderDojo movement. We did so by embedding ourselves in this community and conducting semi-structured interviews. Our analysis of these interviews not only allowed us to shed light on each structural element, it also illustrated several connections between them. It is through active participation in the COP that the participant increases their learning. They learn by doing and being actively involved in the activities of the community. Participation also leads to increased knowledge within the community. Participation in the community activities leads to the dissemination of knowledge as participants communicate, collaborate and share their knowledge with each other. As knowledge is exchanged, learning also takes place. Participation, Learning and Knowledge are interlinked elements of a COP and one cannot take place without the presence of the other two.

As part of our future work, we will plan to carry out further analysis of the interview transcripts to explore whether there is any evidence of other structural elements that have not been established to date. We would also like to expand on the contrasts made between the old and new concepts discussed in this paper, especially in relation to participation and practice. Following this we will focus on developing an approach that will support the creation and development of a successful COP. The aim of this particular COP will be to enhance the learning of STEM subjects

for second level students across Europe. We hope that our new view on how COP are formed and develop will help us and other researchers better understand the wider range of complexities that are involved in the creation, development and study of Communities of Practice.

Acknowledgments

This work is funded by the H2020 project Umi-Sci-Ed[2] [71083]: Exploiting Ubiquitous Computing, Mobile Technology and the Internet of Things to promote Science Education.

References

Amin, Ash & Roberts, Joanne., (2008). Knowing in action: Beyond communities of practice. Research Policy 37 (2008) 353–369.

Avasthi, Vinay., Dey, Shubhamoy., Kamal Kishore, Jain., & Mishra, Rajhans. (2015). The evolution of knowledge in communities of practice. In Proceedings of the 2015 Conference on research in adaptive and convergent systems (RACS). ACM, New York, NY, USA, 96-101.

Bengtsson, Mariette. (2016). How to plan and perform a qualitative study using content analysis. *NursingPlus Open, 2,* 8–14.

Blankenship, Selena. S., Ruona, Wendy E. A. (2007). Professional learning communities and communities of practice: A comparison of models. In F. Nafukho (Ed.), AHRD 2007 International Conference Proceedings, 888-895

Boud, David., Middleton, Heather. (2003). Learning from others at work : communities of practice and informal learning, 15(5), 194–202.

Bradshaw, Pete., Powell, Stephen., Terrell, Ian. (2004). Building a Community of Practice: Group, 184–201.

Brown, John Seely., & Duguid, Paul. (1991). Organizational learning and communities-of- practice: Toward a unified view of working, learning, and innovation. Organization Science 19912:1 , 40-57

Clancey, William J. (1995). A tutorial on situated learning.

Corso, Mariano., & Giacobbe, Andrea. (2005). Building Communities of Practice that work : a case study based research. The Sixth European Conference on Organizational Knowledge, Learning, and Capabilities, 1–20.

Couros, Alec, & Kesten, Cyril. (2003). Communities of Practice: A Literature Review. 1-20.

Eckert, Penelope. (2006). To appear in 2006. Encyclopaedia of language and linguistics. Elsevier. Communities of Practice Penelope Eckert. Communities, 1–4.

Elo, Satu., & Kyngäs, Helvi. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62*(1), 107–115.

[2] See http://umi-sci-ed.eu/

Falkman, Goran., Gustafsson, Marie., Jontell, Mats., Torgersson, Olaf. (2008) SOMWeb: A Semantic Web-Based System for Supporting Collaboration of Distributed Medical Communities of Practice. Eysenbach G, ed. Journal of Medical Internet Research. Res 10(3), e25.

Fischer, Gerhard. (2001). Communities of interest: learning through the interaction of multiple knowledge systems. In: Bjornestad, S., Moe, R., Morch, A., Opdahl, A. (Eds.), Proceedings of the 24th IRIS Conference, Ulvik. Department of Information Science, Bergen, Norway, pp. 1–14.

Gherardi, Silvia. (2006). Organizational Knowledge: The Texture of Workplace Learning. Blackwell Publishing, Oxford.

Graneheim, Ulla. Hallgren., & Lundman, Berit. (2004). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, *24*(2), 105–112.

Helgevold, N., & Moen, V. (2015). The use of flipped classrooms to stimulate students' participation in an academic course in initial teacher education. *Nordic Journal of Digital Literacy*, *2015*(1), 29–42.

Hildreth, P. and Kimble, C. (2002), 'The duality of knowledge', Information Research, Vol. 8 No. 1.

Hoadley, Christopher. (2012). What is a Community of Practice and How Can We Support It? Theoretical Foundations of Learning, (814), 287–300.

Hoadley, Christopher., Kilner, Peter G. 2005. Using technology to transform communities of practice into knowledge-building communities. SIGGROUP Bull. 25, 1 (January 2005), 31-40.

Jakovljevic, Maria., Bushney, Melanie., Buckley, Sheryl. (2013). Forming communities of practice in higher education: a theoretical perspective. International Conference on Management, Knowledge and Learning, 1107–1119.

Kimble, Chris. (2006). Communities of practice: Never knowingly undersold. CEUR Workshop Proceedings, 213, 218–234.

Kimble, Chris., Hildreth, Paul., Wright, Peter. (2001). Communities of Practice : Going Virtual. Knowledge Management and Business Model Innovation, (Sachs 1995), 220–234.

Lave, Jean., Wenger, Etienne. (1991) Situated Learning: Legitimate Peripheral Participation. Cambridge University Press, 1991.

Lindkvist, Lars. 2005. Knowledge communities and knowledge collectivities: a typology of knowledge work in groups. Journal of Management Studies 42 (6), 1189–1210

Nistor, Nicolae., & Fischer, Frank. (2012). Communities of practice in academia: Testing a quantitative model. Learning, Culture and Social Interaction, 1(2), 114–126.

Probst, Gilbert., & Borzillo, Stefano. (2008). Why communities of practice succeed and why they fail. European Management Journal, 26(5), 335–347.

Ramchand, Anand., & Pan, Shan Ling. (2012). The Co-Evolution of Communities of Practice and Knowledge Management in Organizations, 8-23.

Ranmuthugala, Geetha., Plumb, Jennifer., Cunningham, Frances., Georgiou, Andrew., Westbrook, Johanna., Braithwaite, Jeffrey. (2011). How and why are communities of practice established in the healthcare sector? A systematic review of the literature, 1-16.

Roberts, Joanne. (2006). Questioning the place of communities of practice. OLKC Conference at the University of Warwick, (January 2011), 1-21.

Sánchez-Cardona, Israel., Sánchez-Lugo, José., & VŽlez-González, Julia. (2012). Exploring the Potential of Communities of Practice for Learning and Collaboration in a Higher Education Context. Procedia - Social and Behavioral Sciences, 46, 1820–1825.

Sauve, Eric. (2007). Communities of practice. eLearn, 2007(4), 2.

Sheridan, Irene., Goggin, Deirdre., & O'Sullivan, Linda. (2016). Exploration of Learning Gained through CoderDojo Coding Activities. Proceedings of INTED 10th International Technology, Education and Development Conference, 1-8.

Snyder, William. M., & Briggs, Xavier de Souza. (2003). Communities of Practice: A New Tool for Government Managers. Arlington: The IBM Center for The Business of Government, 1-74.

Stein, David. (1998). Situated Learning in Adult Education. ERIC Digest No. 195, 1-7.

Wenger, Etienne. (1998). Community of Practice: a Brief Introduction. Cambridge University Press, 1998, 1-7.

Wenger, Etienne. (1998b). Communities of Practice: Learning, Meaning, and Identity. Cambridge: Cambridge University Press.

Wenger, Etienne. (2006). What are communities of practice? Retrieved from: http://www.ewenger.com/pub/index.htm. Accessed on 12 March 2018.

Wenger, Etienne., McDermott, Richard., & Snyder, William. (2002). Cultivating Communities of Practice. Boston, Mass: Harvard Business School Press.

Wenger, Etienne., & Snyder, William. (2000). Communities of Practice: The Organizational Frontier. Harvard Business Review, 78(1), 139-145.

Wiggberg, Mattias., & Daniels, Mats. (2011). A method for analyzing learning outcomes in project courses. Conferences in Research and Practice in Information Technology Series.

Smit, D.; Lindlbauer, A.; Murer, M.; Hengeveld, B.; Tscheligi, M. (2019): Let the Bot Take Care of It: Exploring #CapIt, a Whiteboard Table Capture System. In: Proceedings of the 17th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing an the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep16

Let the Bot Take Care of It: Exploring #CapIt, a Whiteboard Table Capture System

Dorothé Smit, Andreas Lindlbauer, Martin Murer, Bart Hengeveld* and Manfred Tscheligi
University of Salzburg, Austria; *Eindhoven University of Technology, the Netherlands
dorothe.smit@sbg.ac.at

Abstract. In this paper we describe #CapIt, a system that aims to combine the best of analog whiteboards and digiboards in tabletop collaborations. #CapIt was deployed at an HCI research unit for over a year. In an exploratory study completed after the system was fully integrated in participants' day-to-day work practices, three power users of #CapIt were asked to reflect on their use of the system by means of mindmapping. Using photo elicitation and semi-structured interviews, additional feedback was gathered from the participants. Based on the participants' comments as well as our observations of the mindmapping process, we here report our findings pertaining to (1) hybridity; (2) collaboration; and (3) territories, privacy and temporality; and discuss the influence of the system on collaborative work practices.

1 Introduction

Whiteboards are an often used tool in nearly all types of work, from individual whiteboards in personal offices, to public whiteboards in meeting rooms, to the often digital whiteboards in the front of classrooms. There are several advantages of analog whiteboards that keep attracting people to pick up a marker and draw out their thoughts: whiteboards never have to reboot and never suffer technical failures

(Price et al., 2011), there is no learning curve for using a whiteboard (Gumienny, 2013, p. III), and quick whiteboard sketches support problem solving and reasoning (Larkin and Simon, 1987). As communal whiteboards continue to be frequently used in many workplaces, it is beneficial to research how to best capture, digitize, and share notes from a whiteboard, so that collaborators can easily integrate them into their work practices.

In this exploratory paper, we describe #CapIt, a system that takes snapshots of the notes and sketches on a whiteboard that also functions as a table (see figure 1). The system uploads the captures to the internal messaging system of an HCI research unit. The whiteboard table and its capture system are used regularly for, e.g., collaborative ideation sessions, team meetings, and project progress updates. We illustrate how the characteristics of an analog whiteboard, paired with the affordances of a table, and combined with an already familiar messaging and archiving system, influenced collaborative and creative meetings at the research unit. The findings described in this paper relate to (1) the hybridity of the system in terms of digital vs. analog, but also horizontal vs. vertical; (2) the effect of the system on collaboration, relating to the type, significance and sharing of notes; and (3) a category that encompasses three interlocking themes: territories, privacy and temporality.

Figure 1. Left:The Whiteboard Table, as it is placed in the Center for HCI, spans 600×120 cm, and is surrounded by $\sim$ 12 chairs at any given time. Right: a generic team meeting around the whiteboard table. © David Fisslthaler.

2 Background

To frame our research, as well as the #CapIt system, we will briefly discuss existing research about (collaborative) work on (electronic) whiteboards, as well as research about large, communal table tops.

In many work domains, the use of whiteboards continues to be universal. In the fields of design, architecture, engineering, and computer science, sketching and diagramming is a well-established practice (Walny et al., 2011). In this section, we will discuss the theory behind the use of (digital and analog) whiteboards and

whiteboard capture systems, of which there are many. We will also discuss the characteristics of tabletop collaboration, and how territoriality in shared work spaces influences work practices.

2.1 Whiteboard Note Taking

Content on whiteboards helps externalize thoughts, so they can be more easily understood by others, and supports exploring ideas, without taking decisions too quickly (Cherubini et al., 2007; Triplett, 2016). Whiteboards are used both independently or collaboratively, and synchronously or asynchronously (Tang et al., 2009). In asynchronous use, users of the whiteboard leave notes, sketches and diagrams behind for themselves or others to work on (Mangano et al., 2015). Content on a whiteboard is easily revisitable, updatable, and flexible, allowing users to build representation of information for many types of collaborative and individual activity (Tang et al., 2009).

2.2 Electronic Whiteboards

Electronic whiteboards are popular due to their capability of combining the properties of an analog whiteboard with other (digital) teaching tools (e.g., showing videos or 'undoing' and 'redoing' steps). However, electronic whiteboards can lack resolution (Branham et al., 2010), and they are often turned off to preserve energy (Huang et al., 2006), which causes them not to be *'ready-to-hand'* (cf. (Heyer and Brereton, 2010)), a requirement for skilful, flowing use. Analog whiteboards are therefore still often used by researchers, designers and engineers to understand their own work, as well as communicate it to others.

2.3 Sharing Whiteboards

Communally shared analog whiteboards face the problem of ownership, i.e.: sketches and notes created on a public whiteboard are at risk of being wiped out by an external party (Price et al., 2011), causing people to write notes like 'do not erase!' on whiteboards (Saund, 1999). This results in some notes remaining on the whiteboard for very long times, due to externals' fear of removing important work, and thus rendering the whiteboard useless (Ju et al., 2007). Additionally, it is hard to digitize the notes for archiving purposes: rather than copy the notes by hand by means of a text processor, whiteboard users can often be seen taking cell phone pictures of the whiteboard, for future reference (Inie and Dalsgaard, 2017). Photographs of whiteboards are often used to solve disagreements and to confirm action points (Walny et al., 2011), but to effectively do this, the photographs need to be shared with the entire group involved in the meeting by the person who took them.

Due to practical considerations, it is also rare to see more than two persons writing on the same whiteboard at the same time, as space generally does not allow

for it. When multiple people write on a vertical whiteboard, the overview and communal understanding of what is happening on the whiteboard is lost quickly.

2.4 Capturing Whiteboards

As common as whiteboards are in office environments, so are smart phone snapshots that try to preserve the outcomes a creative, collaborative session on a shared whiteboard (Klokmose and Bertelsen, 2013; Branham et al., 2010).

Varona-Marin et al. (2018) recently analysed the curation of manually captured photos of the whiteboard after the meeting has ended. They found that snapshots of the whiteboard usually serve as general meeting records. Even though only a single group member usually takes a photo of the board, the captures were perceived by the users to be collectively owned by the group members. The captures were often shared via email, or stored in shared folders.

In general, captures of whiteboards seem most often revisited when either participants in a meeting differ on a decision made during a meeting, or when sketches, notes and diagrams made during a meeting need to be digitized to be used in further work (such as reports and presentations) (Walny et al., 2011).

2.5 Working on Large Table Tops

Large tabletops invite more explorative or playful interactions with objects on a tabletop (Zagermann et al., 2016). Tabletop collaboration increases the awareness of the actions of other participants (Rick et al., 2011); it equalizes the roles of the participants (Marshall et al., 2008); it encourages more cohesive work (Rogers and Lindley, 2004); and resolves bottle necks (Tang, 1991). A key goal of collaborative work is often collaborative sensemaking: bridging gaps in understanding between people (Wallace et al., 2013). People working on complex projects tend to externalize key aspects of their sensemaking process, to literally and physically 'lay out the evidence' on table (Andrews et al., 2010).

2.5.1 Tangible Objects

Large spaces support more explorative interactions, involving fidgeting or playing (Zagermann et al., 2016). Wall-mounted whiteboards – both digital and analog alike – offer little space for tangible objects to come into play. Artifacts like paper, pens and other peripherals are primary tools for explaining, developing and communicating ideas during early phases of design (Klemmer et al., 2001). These kinds of objects often act as placeholders in the early stages of design (Smit et al., 2016), functioning as *scaffolds* (Jaasma et al., 2017) or *traces* (van Dijk and Vos, 2011), to support the designers in their process of exploring, extrapolating and communicating. Large, horizontal surfaces more practically allow for interaction with physical, three-dimensional objects, which can support collaborative sensemaking processes (Hummels and van Dijk, 2015). The table surface is an important resource for collaboration mediation; and the spatial orientation of the

participants in relation to each other and the drawing plays a role in the structure of the activity (Tang, 1991).

2.5.2 Orientation

Orientation is critical to how people comprehend information, coordinate actions with one-another, and mediate communication. Orientation plays a major role in informing collaborators who is currently using or reviewing which items, and which items are available. Collaborators often rotate items on a table partially (i.e., sideways) to share the item with others and invite immediate collaboration (Kruger et al., 2004).

2.6 Territoriality in Shared Workspaces

Collaborative work on a shared surface also introduces with so-called *territoriality*: tabletop territories serve to coordinate tabletop interactions (Scott et al., 2004). Territories help people coordinate tasks and create mutual understanding, and so their establishment is crucial in the beginning of a collaborative task (Klinkhammer et al., 2018). Collaborators around a table automatically define personal territories for themselves, in which they collect items and do work that relates only to them. Although never explicitly discussed, collaborators hardly ever venture into another person's personal territory (Scott et al., 2004). Participants may even ask for explicit permission to add to, or adjust items in another person's territory, even if those items are not personal (Morris et al., 2010).

If no personal territories are established, conflicts may arise (e.g., because participants interact with materials that 'belong' to another person) (Pinelle et al., 2009). Personal territories generally reside along the edge of the table, in front of the respective participant (Klinkhammer et al., 2018), group territories take up the remaining space on the table (Scott et al., 2004). Workspace territories are not static states, but instead change shape following the flow of the collaborative process (Klinkhammer et al., 2018).

2.7 Previous Whiteboard Note Capture Systems

In the following sections, we will describe previous works that involve whiteboard note capture systems. Each of the works discussed presented findings that relate to the use of our #CapIt system.

An early adoption of a whiteboard capture system is the ZombieBoard (Saund, 1999). This work featured a pan/tilt camera that would construct a high-resolution capture of a whiteboard by *mosaicing* several pictures together seamlessly. The capture was then automatically printed. The researchers found that a privacy blind, installed in front of the system's camera to obstruct the camera's view, was occasionally used.

Zhang and He (2003; 2004) describe a system for scanning whiteboard content by means of a digital camera. In this system, entire meetings were captured on

video. Even when use of the whiteboard was not required for the meeting, people still turned on the capture system, mainly to capture who was speaking (by writing the name of the speaker on the whiteboard) so that meeting segments could be recognized and retrieved more easily later on.

Holmquist et al., (2003) introduce Total Recall, in which the user holds a hand-held computer with screen up to the board and moves it around to recover previous notes taken on that area of the whiteboard. They argue that this solution provides a better coupling to the whiteboard notes than viewing a capture of the whiteboard on a desktop system.

Price et al. (2011) used wireless-enabled digital cameras to take pictures of students' personal whiteboards, that they used in class. Those photos were then uploaded to a photosharing website. They found that students would diligently label and organize their whiteboard pictures for later use. Additionally, students began to correct the solutions on their personal whiteboards before capturing them, ensuring that a capture showing the correct solution to a problem was uploaded.

Branham et al. (2010) describe ReBoard, a system that focuses enabling detailed search within the collection of whiteboard captures. The search function of ReBoard was based on general date ranges, thumbnails and general location of a sketch on the whiteboard. They found that users shared captures either through the system or via personal email, or that the images were sometimes printed to share with others.

3 Research Objective

In this exploratory paper, we describe the use of the whiteboard table in combination with the connected capture system, that automatically uploads whiteboard captures to the messaging system used in the workplace this exploratory study took place in. The system that we present, #CapIt, combines the advantages of the horizontal orientation of the whiteboard table, the ease-of-use of analog whiteboards, and the archiving capabilities of digiboards. Therefore, our research focuses on the use of the whiteboard table capture system in collaborative settings, and the advantages of the table's horizontal orientation and the system's connection to existing digital infrastructure.

4 System

At the Center for Human-Computer Interaction in Salzburg, a multidisciplinary team of 30 researchers investigate HCI problems. One of the tools they use for this, is the whiteboard table: a structure of 600×120 cm, comprised of two horizontal whiteboards (300×120 cm each; see figure 1). The researchers of the Center use this table on a daily basis, both for collaborative, as well as individual

work. The researchers of the Center also use Slack[1] as their main means of digital communication. Slack offers public and private chatrooms called 'channels' (e.g., *#channel*) as well as a direct messaging system for one-on-one and group conversations. Files, such as photos, can also be shared in Slack. Users of Slack are encouraged to create their own 'bots' to automate processes for them. Bots can do many things that human users of Slack can also do, such as sending messages and sharing files.

To combine the unifying power of the whiteboard table with the easy sharing of files and messages in Slack, #CapIt was created. #CapIt is a system that captures notes and sketches made on the whiteboard table with the press of a button. The captures would then be uploaded in Slack by the *@whiteboardbot* to the *#whiteboardchannel* (see figure 2), where users could review, download and share the photos that were taken of their notes.

As early as 1988, research showed that one major reason that groupware (interactive software and hardware in the workplace) fail, is because they ask more time and energy from the users, than they are getting in return (Grudin, 1988), and still, groupware often fails be the systems are too complex or badly designed, and it's easier for the user to avoid using them altogether (Korpelainen and Kira, 2013). #CapIt was therefore designed with simplicity of use in mind, integrating the capture system fully with the messaging system that the future users of #CapIt were already using. The system is designed in a rather open-ended fashion, so that users may interpret the system in ways that we, as the designers of the system, could not have fully foreseen (Pipek and Wulf, 2009), in hopes that the users appropriate the system in a way that most effectively and efficiently supports their work day.

Figure 2. The *#whiteboardchannel* in the Slack work space, where the captures of the whiteboard table are uploaded for the researchers of the Center for HCI to use.

[1] https://slack.com/

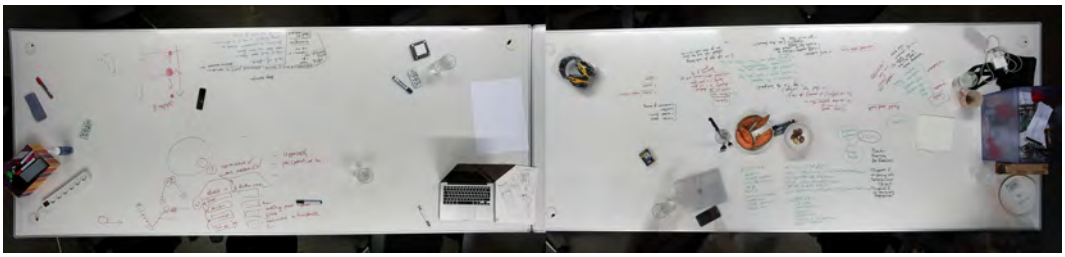


Figure 3. A typical picture as taken by #CapIt. In this picture, one can see the type of notes taken during meetings and creative sessions, but also the presence of other, tangible items, including: basic whiteboard table necessities (markers, wipers); technology (personal computers, phones, chargers); and peripherals (snacks, water glasses, coffee cups).

4.1 Technology

To capture the entire length of the whiteboard table with a high enough resolution to maintain legibility of small, handwritten notes, we needed a minimum of 25 points per inch[2] (PPI) (Zhang and He, 2004). The PPI can be calculated by dividing the diagonal resolution of a picture by the diagonal size of the subject (in this case: the whiteboard table) in inches (in this case: 241"). This leaves us with a minimal resolution of 1182×5908 pixels – more than 4K resolution. Therefore, the final iteration of #CapIt makes use of *two* 4K cameras, the images of which are stitched together. These two cameras are mounted 2 meters above the whiteboard, and connected to a ODROID-XU4 single board computer[3], by means of two USB 3.0 ports. The cameras are triggered by a Logitech POP button[4] (see figure 4). This button sends a signal via Bluetooth Low Energy (BLE) to the ODROID board. Once the pictures have been taken, they are stitched together (see figure 3) and uploaded to the public *#whiteboardtable* Slack channel, where users can find, download and share the pictures of the whiteboard table.

Figure 4. The cameras above the whiteboard table are triggered by this white Logitech POP button, which was placed in a black, laser-cut casing (whiteboard markers for scale).

[2] 1 inch (1") $\approx 2.54cm$

[3] https://www.hardkernel.com/shop/odroid-xu4/

[4] https://www.logitech.com/en-us/product/pop-smart-button

5 Study Setup

#CapIt has been in place at the Center for HCI for more than a year, during which employees have freely been making use of the system. All pictures (n = 168) that were shared in the *#whiteboardtable* channel in Slack during the period of 21 November 2017 to 21 November 2018, were downloaded. 79 Pictures were discarded for being outside the scope of research (e.g., duplicates or pictures that were taken in quick succession after one-another, pictures that did not show any notes, or pictures that were taken during prototyping and testing phases). This resulted in a data set of 89 pictures.

5.1 Participant Selection

From the dataset, in which 17 different users of the system appeared, we identified three so-called 'power users': the users that appeared most often in the pictures. For this study, the three power users appeared in 18, 13, and 12 pictures respectively. We identified users not only based on appearance (i.e., clothing, posture), but also on personal items, such as: stickers on laptops, headphones, handwriting in the notes, personal water bottles, and notebooks, which were visible in the photos. In 25 pictures, no user could be indubitably discerned.

5.2 Generative Tools

During the study, we followed an approach developed by Keller (2005, p. 23-27), who used techniques from the field of participatory design to elicit responses from experts about their methods of collecting and structuring inspirational material. Keller (2005) employed generative tools (2000). Specifically, participants were asked to create three mindmaps relating to the way the participants structure their collections of creative and inspirational material.

The three power users were invited for a collaborative mindmapping session that served to learn about their use of the whiteboard table documenting system. By allowing for visual expression of the participants in the mindmaps, rather than just verbal expression in a semi-structured interview, we hoped for more diverse insights shared by the participants (Keller, 2005, p. 23). Different from Keller's method, we opted for a collaborative session, as #CapIt is mostly used in collaborative sessions. Participants, therefore, are used to sharing the space on the whiteboard table with others, and collaboratively create notes.

By inviting the participants to *use* the system, with the goal of exploring and studying the use of the system, we hoped to inspire the participants to share recursive feedback, in a dialogue not just with the researchers, but also with the system.

5.3 Participant Driven Photo-Elicitation

To start a dialogue about the effect of #CapIt on the work of the participants, the researchers opted to begin the interview section of the study with a participant driven photo-elicitation session (Harper, 2002; Van House, 2006; Gorm and Shklovski, 2017). Photo elicitation has been shown to help participants focus on the interview, and to make new associations (Carter and Mankoff, 2005). Before the study, the participants were asked to send the researchers the most interesting picture they had taken with #CapIt. The researchers then edited these pictures to remove anything that wasn‘t a note or sketch (see figure 5). The participants were asked to explain what stood out to them about the edited photos. Then, the unedited photos were shown to the participants, and they were asked to compare the pictures and comment on the differences.

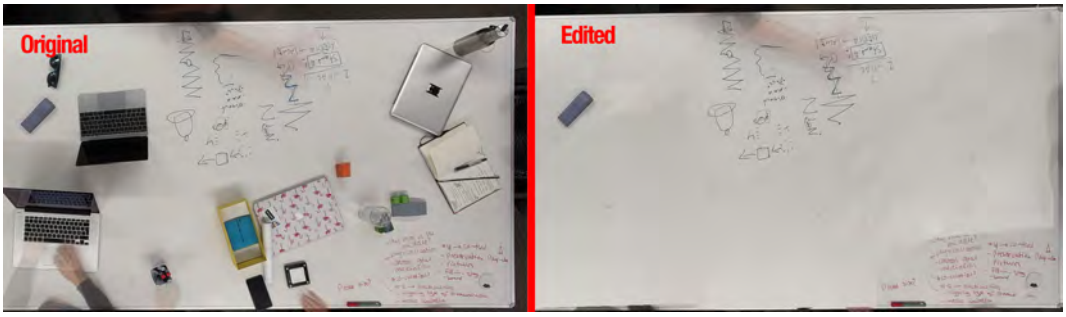


Figure 5. The pictures that the participants had sent to the researchers were edited to remove anything that would not be on a ‘normal’ (vertical) whiteboard. Left: part of the original picture that P1 sent in. Right: part of the edited picture.

5.4 Semi-Structured Interview

Directly following the photo-elicitation process, the researchers transitioned into a semi-structured interview that touched on the use of the whiteboard table; the capture system; and the sharing, archiving and sensemaking of the photos. Although questions were prepared for the interview, they were not asked sequentially, but rather guided the interview while keeping the flow of the conversation intact.

6 Data Collection & Analysis

Two researchers involved the development of #CapIt were present during the entire study. The primary researcher guided the mindmapping, photo-elicitation and interview sessions, while taking notes. The secondary researcher was responsible for monitoring the recording equipment (a GoPro Hero 6 recorded the entire session, while the semi-structured interview was additionally recorded on a smartphone), and took notes as well. #CapIt itself was used to intermittently capture progress of the mindmapping on the table.

The notes of both researchers were digitized and merged in one document. These notes included observations of the participants, Both researchers then, independently, engaged in descriptive coding of the study notes (Saldana, 2009, p. 6). These two separate sets of codes were then compared, and the notes were categorized based on the codings. From these categories, ten overarching, often interconnected themes were derived. They are described in the findings.

The mindmaps that were made by the participants during the study were captured by #CapIt (see figure 6). While the specific content of the mindmaps themselves were not extensively archived, we analyzed the process of making them, and their outcomes, based on #CapIt captures and research study notes.

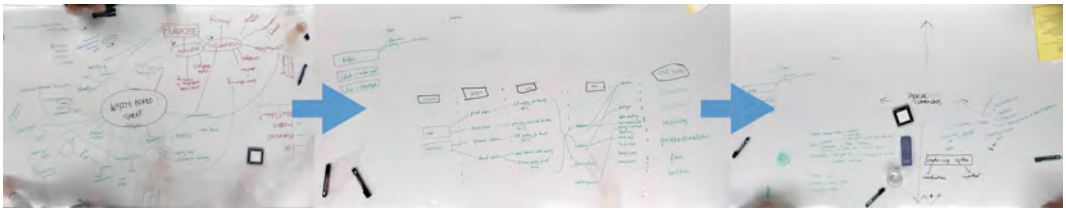


Figure 6. The three mindmaps that were created by the participants during the study. From left to right: 1) the first mindmap about the general use of the whiteboard; 2) an attempt at structuring the initial mindmap in terms of users, purpose, role, use, and end goal; 3) a more specific mindmap about the physicality of the whiteboard table, and the accompanying capturing system.

7 Findings

In this section, we will describe the most interesting findings from the study by means of describing the overarching themes that were found during the analysis of the study results.

7.1 Choice of Picture

An immediate interesting finding is illustrated by the choice of pictures of the participants. Only P1 submitted a 'standard' #CapIt picture; i.e., the type of picture that is most commonly captured by #CapIt, including meeting notes and to-do-lists. She mentioned that she had chosen this note as an example of how she would refer back to a #CapIt picture, when she had not taken notes in her personal notebook during a meeting.

P2, on the other hand, submitted a picture that was taken during a creative workshop with children in the context of a research project. She mentioned that she not only took the picture as documentation of what had taken place during the workshop, but that the picture also perfectly served the purpose of sharing their progress in external presentations, as faces are not visible in the picture.

Finally, P3 submitted a very old picture, that she had saved somewhere in her personal files, and named *'Data_Analysis_Observations.jpg'*. She clarified that this

was the picture that she most often referred back to, as it contained the entire data analysis of a project she was working on. She mentioned that she had even forgone making a spreadsheet on her computer, but had instead saved this capture to the Data Analysis folder on her computer. According to P3, the collaborative process of data analysis on the whiteboard table started quite chaotically, but by being able to all work on it collaboratively, and by being able to erase and rearrange items easily, it became a structured overview.

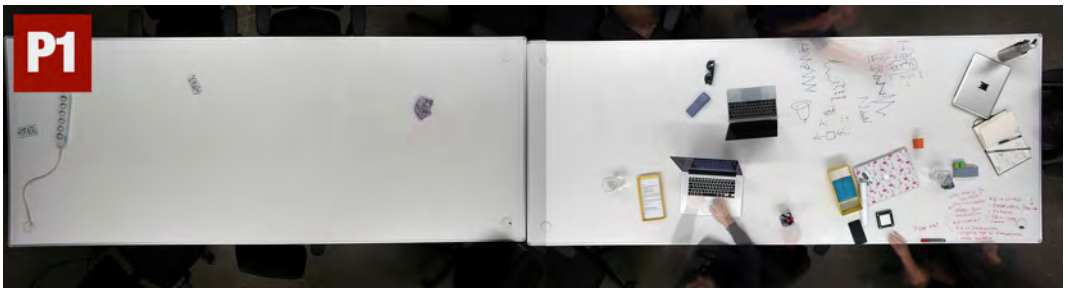


Figure 7. P1 submitted a picture of a recent meeting she took part in. Although she did not write any of the notes on the table herself, she explained that this picture was important to her, because she did not take any private notes in her notebook during the meeting, and went back to it several times to recall what steps were agreed on, and what her tasks were.

Figure 8. P2 chose a picture taken during a creative workshop with children, which took place in the context of a research project. She chose this picture, not so much for the content in it, but for the message that is conveyed with it. She has used the picture several times in presentations for external parties, to explain the kind of work she does within this project.

7.2 Type and Significance of Notes

The participants were asked their first thought towards 'working on the whiteboard-table'. P1 (designer) mostly associated the whiteboard table with drawing, while P2 (designer) generally used in to explore thoughts, both in groups and individually. P3 (sociologist) said to mostly associate the whiteboard table with data analysis. The responses of the participants show that #CapIt is useful for a range of different activities, and for people with different backgrounds. Captures reviewed for this study included not only meeting notes and brainstorming sessions, but workshop progress, physical prototyping sessions, artistic drawings, data analysis, time lines and schedules, etc.

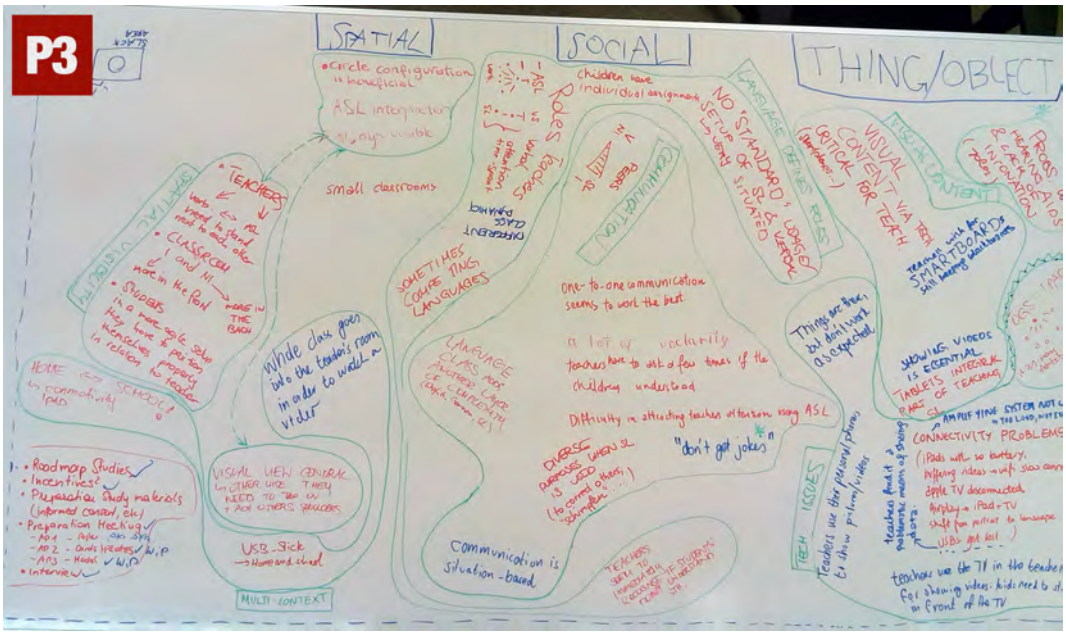


Figure 9. P3 found a picture that she had saved to her personal collection on her computer for future reference: a photo of the whiteboard table after an intensive collaborative data visualization session. Note: this picture was taken with an older version of the system, hence the lower quality image.

Participants noted that when they come across notes from others on the table, they can generally deduce significance (P2: *"From the way the notes are written down, I can decide if they're important"*). They mentioned that they would generally not delete things that had colleagues' names attached to them, or things that were appropriately titled (P2: *"Stuff with titles like 'Data Analysis', I would never delete"*), unless they absolutely needed the space to work on. In those cases, P2 and P3 would capture the work before erasing it, while P1 said that she would then rather not work on the whiteboard table.

Although the participants mention that they think they can generally deduce the importance of notes and sketches from their appearance, P3 also reported on an instance where a colleague drew a simple scribble during a project meeting to illustrate an epistemological point (see figure 10). Although this colleague never meant for the sketch to be saved, it became a guiding visual keepsake for P3, to remind her of the main objective of the project.

7.3 Roles & Collaboration

Right from the start of the mindmapping session, dividing roles was noted as an important step in the use of the whiteboard table (P3: *"Maybe we should divide colors* [of markers]*?"*), especially after it became clear that a lot of terms appeared several times in the mindmap (P3: *"We are reproducing a lot of stuff: we need roles."*, and "[In a normal session] *we would have decided: who will write?*"). The participants mentioned that in meetings with a collaborative goal, there is generally

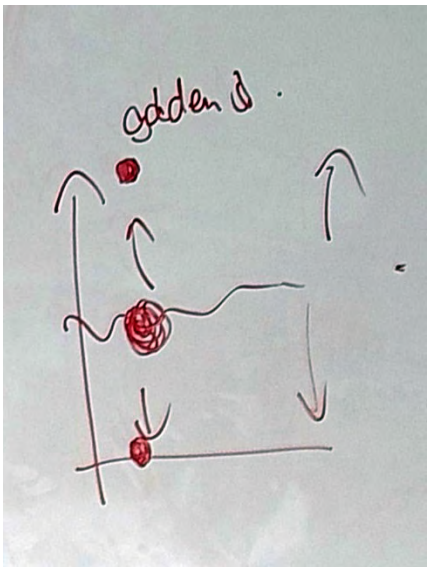


Figure 10. A seemingly innocuous scribble, that came to play an important role in the stucturing of a complex project.

one person taking on the main writing task, with others adding on where needed. In meetings during which people have their individual tasks and goals people tend to write their own, personal notes, in their personal territory.

7.4 Structure

During the study, the participants started out with an unstructured brainstorm (see figure 6, left), calling out and writing down whatever came to mind when thinking of #CapIt. Nonetheless, the first mindmap ended up having multiple entries of the same term. Realizing that they could not reach sufficient structure in the current mindmap, they decided to create a new mindmap on the other side of the table. Whereas the first mindmap had the characteristics of a word web, the second mindmap took the shape of a flowchart or schematic. The participants found themes in the original mindmap, and structured them under the headers of: users, purpose, role, use, and end goal (see figure 6, middle). Finally, the third mindmap resembles a chart, in which the participants tried to map characteristics of the system to physical versus digital components (see figure 6, right).

7.5 Orientation

As any horizontal shared workspace, the whiteboard table faces issues of orientation of content. P3 even mentioned that she has become very skilled at writing upside-down. The participants noticed while making the first mindmap, that a lot of the duplicate terms were written in different orientations (i.e., the participants might have noticed the term if it had been written in the 'correct' orientation). This resulted in the decision of the participants to sit on the same side

of the table while making the second and the third mindmap (P3: "*If we're on the same side, it's easier to share note-taking* [tasks]". Furthermore, the participants switched seats twice while making the first mindmap, to have a look at the mindmap from the other participants' perspectives.

7.6 Effects of Removing Peripherals

Once the peripherals were removed, the difference in the type of pictures that the participants had brought along for the study, became all the more clear. Whereas the focus of P3's picture is fully on the notes that are written on the table, the picture that P2 has brought completely loses all meaning when the tangible workshop materials are removed. In this picture, the tangible materials on the table play a bigger role in the activity going on around the table, than the sketches on the table themselves.

In the picture that P1 brought, removing the tangible objects on the table did not change the meaning of the notes, however: the ownership of the notes (i.e., the personal territories of the participants) could no longer be identified, due to lack of personal items on the table. Additionally, the main subject of the meeting – a yellow box, seen in the middle of the right table in figure 7 – was also removed, making it more difficult to recognize the capture at a glance without reading the notes in the image.

7.7 Trust in the System

P3 noted that whenever she takes a picture of her work on the whiteboard table, she checks the *#whiteboardtable* channel in Slack to see if the picture is there, before erasing her work. The other two participants noted that they did not do that: they trust the system. We suspect this to be a result of P3's use of a prototypical version of #CapIt. The reliability issues of this version of #CapIt may have led P3 to believe that the system lacks functionality or predictability, and that the system therefore cannot be trusted (Thatcher et al., 2011). Rather than discontinuing the use of the system, she instead decided to always check that the system behaved according to her intentions. P1 and P2 only every experienced a fully functional system, and therefore did not experience any disconnect between system expectation and system confirmation (Bhattacherjee, 2001) and did not feel the need to double-check on the system's performance.

7.8 Temporality vs. Permanence

The system causes tension between temporality and permanence. All participants viewed notes on the whiteboard as a work-in-progress (P2: *"It is kind of like real-time editing"*; P1: *"I can always take it back"*): participants felt less need to prepare or structure their thoughts before writing them down, than for paper note-taking – which is shown to stimulate creativity (Diehl and Stroebe, 1987). However, the addition of #CapIt with its automatic upload to a public Slack channel, introduces a layer of permanence to scribbles and sketches (P2: *"You cannot press a button*

to delete a photo. Once it's taken, it's out there"). Capturing the notes therefore takes a way the "erasability" that is a defining quality of whiteboards. P3 noted that she often cleans up her notes and sketches, and removes peripherals from the table, before pressing the capture button (cf. Price et al. (2011)).

7.9 Territory

As in other work, we also noticed that personal territories are quickly created on the whiteboard table. All participants shared the habit of writing small, personal notes (such as to-do lists or reminders) on the edge of the table, whereas communal note-taking happens collaboratively in the middle of the table. The personal territory on the whiteboard was classified by the participants as being somewhere between a personal notebook (completely private), and the center of the table (P3: *"It's a big table, you're not always in control of what people write and especially what they erase"*).

7.10 Privacy Concerns

Mounting cameras in office spaces always come with concern for individuals' privacy (Saund, 1999; Branham et al., 2010). Combined with an automatic uploading system, that shares the image in a channel that is accessible to approximately 30 people, the issue of privacy regularly came up in the interview with the power users. All participants agreed however, that anything written on the whiteboard table and not immediately erased, should be considered public knowledge (P3: "*If people leave stuff up for days, it can't be NDA anymore*"). When P1 mentioned that she would not feel comfortable taking pictures of other people's notes that have been left on the table, P2 and P3 both affirmed that they would take captures of other peoples' notes before erasing them, to ensure that they did not delete any important work.

8 Discussion

In this section, we will discuss implications of the findings from the study, as well as the significance of some of the researcher's most interesting observations. The themes that were described in the findings, and the relations between them, will be discussed below.

8.1 Hybridity

#CapIt introduces a hybridity that has, to our knowledge, not yet been described in literature. Capture systems for vertical whiteboards exist aplenty (see, e.g., Fakih (2012); He et al. (2003); Zhang and He (2004); Varona-Marin et al. (2018)), as do interactive whiteboards (see, e.g., Saund (1999); Rebecca et al. (2015). On the horizontal front, there are many examples of interactive tables (e.g., Wallace et al.

(2013); Zagermann et al. (2016); Rogers and Lindley (2004); Rick et al. (2011). #CapIt, however, combines the advantages of an analog whiteboard (foolproofness), with the advantages of a digital whiteboard (archiving), and those of horizontal surfaces (face-to-face collaboration and interaction with tangible objects).

These characteristics lead to unique interactions between the physical and the digital. For example, we expected that participants would depend on the capture date of the picture and the (written) notes (Branham et al., 2010), to discern which picture was of importance to them, but we actually noticed that **users of the system use physical peripherals (personal notebooks, smart phones, even people's hair colours and clothing) to recognize pictures**. Removing the tangible objects from the picture, such as can be seen in figure 5, also removes the context of the picture to a large extent.

Furthermore, we often noticed that **physical objects were often part of the activities taking place on the whiteboard table**, as can be seen in figure 8, where the tangible objects play the main role in the picture, and in figure 7, where a tangible object (the yellow box) provides the context of the picture, and the meeting that took place.

That tangible objects are important for creative, collaborative work is not new information (van Dijk and Vos, 2011). However, we found that **some qualities of these tangible objects appear to also translate to, and even enhance, purely digital content in a different context** (i.e., on a computer or screen, *after* the meeting).

8.2 Collaboration using #CapIt

We found that the whiteboard table in conjunction with #CapIt invites many different modes of use, for many different types of collaborative work. **For collaborative work that takes place in the shared territory of the whiteboard, the orientation of notes, sketches and physical objects is very important**, whereas the orientation of personal notes in the personal territory only matters to the person taking the note. The system also seems to increase the feeling of control over the shared territory of the whiteboard table, because although anyone can add and erase notes, everyone is also free to take a snapshot, so that nothing is lost.

The finding that importance of notes was not always immediately acknowledged, e.g. Tang et al. (2009): the meaning of sketches and notes can evolve over time, transforming from situated, contextual drawings into personal reminders, communication aides, brainstorm starting points, etc. The direct connection between the whiteboard table and Slack becomes important here: participants may not normally have taken the time to copy down the notes and sketches in their own notebooks, but **it is hardly any effort to hit the capture button at the end of a meeting**. Even if participants do not believe that there is anything worthwhile in the notes, they will still have it archived in the *#whiteboardtable* channel, just in case. #CapIt accordingly provides a press-of-the-button back-up system for its users that is readily available to all

participants at all times. We believe that this **continuous availability and the ease of use are a large part of the success of #CapIt** and any other type of collaborative work support system.

8.3 Territories, Privacy, and Temporality

The themes of privacy and territory also strongly interacted with the theme of temporality versus permanence. Whereas others describe that whiteboard capture system often lead to privacy concerns regarding the content of meetings (Saund, 1999; Branham et al., 2010), the participants in this study were **primarily concerned with their personal territories, and the interplay between a very temporary note on a whiteboard, and a very permanent capture**, that is public for everyone to see. The physical capture button – combined with the sound effects the system makes – eases some concerns: it is not possible to covertly take a snapshot of the contents of the whiteboard table.

Furthermore, one participant reported that she would not take pictures of others' work, which is a surprising finding, as there is no indicator as to who took a picture: **the ownership of the picture is shared between anyone who can access the *#whiteboardtable* channel**. However, for some users, the territory of the person creating notes and sketches on the whiteboard table remains with those notes, until they are removed by the person who created them.

#CapIt was not subjected to the strong feelings about privacy that other systems have been. This is likely related to the fact that the system was implemented in an open office structure, where the expectation of privacy is already low. It may also be a sign of developing times, in which privacy cannot be assumed, with or without a system like #CapIt present. In any case **the acceptance of the system in terms of privacy may be heavily dependent on the spatial context of the system**.

9 Conclusion

In this paper, we have described an exploratory study regarding a whiteboard table capture system, that took place at an HCI research unit. Three power users were asked to make mindmaps using the system, followed by a photo elicitation session with captures from the system, and a semi-structured interview. Ten themes were defined in the findings, that connected into three categories that were described in the discussion: the hybridity of the system; the effect of the system on collaboration; and a category that encompasses three interlocking themes, namely: territories, privacy and temporality. #CapIt has been organically embedded into the work practices of the researchers at the HCI research unit, and the captures often find their way out of the *#whiteboardtable* channel into the day-to-day business of the users of the system, even (long) after meetings around the whiteboard table have concluded.

References

Andrews, C., A. Endert, and C. North (2010): 'Space to think: large high-resolution displays for sensemaking'. In: *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. Atlanta, Georgia, USA, p. 55, ACM Press.

Bhattacherjee, A. (2001): 'Understanding Information Systems Continuance: An Expectation-Confirmation Model'. *MIS Quarterly*, vol. 25, no. 3, pp. 351.

Branham, S., G. Golovchinsky, S. Carter, and J. T. Biehl (2010): 'Let's Go from the Whiteboard: Supporting Transitions in Work Through Whiteboard Capture and Reuse'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 75–84, ACM.

Bruun, A., K. E. Jensen, D. H. Kristensen, and J. Kjeldskov (2017): 'Escaping the Trough: Towards Real-World Impact of Tabletop Research'. *International Journal of HumanâĂŞComputer Interaction*, vol. 33, no. 2, pp. 77–93.

Buzan, T. and B. Buzan (1996): *The mind map book: How to use radiant thinking to maximize your brain's untapped potential*. Plume New York.

Carter, S. and J. Mankoff (2005): 'When participants do the capturing: The role of media in diary studies'. In: *In Proceedings of the Conference on Human Factors in Computing Systems (CHI*. pp. 899–908, ACM Press.

Cherubini, M., G. Venolia, R. DeLine, and A. J. Ko (2007): 'Let's Go to the Whiteboard: How and Why Software Developers Use Drawings'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 557–566, ACM.

Cook, D. J. and B. P. Bailey (2005): 'Designers' Use of Paper and the Implications for Informal Tools', pp. 10.

Dickson, P. E., W. R. Adrion, and A. R. Hanson (2008): 'Whiteboard Content Extraction and Analysis for the Classroom Environment'. In: *2008 Tenth IEEE International Symposium on Multimedia*. Berkeley, CA, USA, pp. 702–707, IEEE.

Diehl, M. and W. Stroebe (1987): 'Productivity loss in brainstorming groups: Toward the solution of a riddle.'. *Journal of personality and social psychology*, vol. 53, no. 3, pp. 497.

Fakih, A. (2012): 'Real-Time Whiteboard Capture System'.

Glover, D., D. Miller, D. Averis, and V. Door (2005): 'The interactive whiteboard: a literature survey'. *Technology, Pedagogy and Education*, vol. 14, no. 2, pp. 155–170.

Gorm, N. and I. Shklovski (2017): 'Participant Driven Photo Elicitation for Understanding Activity Tracking: Benefits and Limitations'. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17*. Portland, Oregon, USA, pp. 1350–1361, ACM Press.

Grudin, J. (1988): 'Why CSCW Applications Fail: Problems in the Design and Evaluationof Organizational Interfaces'. In: *Proceedings of the 1988 ACM Conference on Computer-supported Cooperative Work*. New York, NY, USA, pp. 85–93, ACM.

Grudin, J. and L. Palen (1995): 'Why Groupware Succeeds: Discretion or Mandate?'. In: H. Marmolin, Y. Sundblad, and K. Schmidt (eds.): *Proceedings of the Fourth European Conference on Computer-Supported Cooperative Work ECSCW âĂŹ95*. Dordrecht: Springer Netherlands, pp. 263–278.

Gumienny, R. (2013): 'Understanding the adoption of digital whiteboard systems for collaborative design work', pp. 154.

Harper, D. (2002): 'Talking about pictures: A case for photo elicitation'. *Visual Studies*, vol. 17, no. 1, pp. 13–26.

He, L.-w., Z. Liu, and Z. Zhang (2003): 'Why take notes? Use the whiteboard capture system'. In: *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03)*, Vol. 5. pp. V–776–9 vol.5.

Heyer, C. and M. Brereton (2010): 'Design from the Everyday: Continuously evolving, embedded exploratory prototypes'. In: *Proceedings of the 8th ACM Conference on Designing Interactive Systems: August 16-20, 2010, Aarhus, Denmark*. New York, N.Y., ACM Press. OCLC: 669871371.

Higgins, S., G. Beauchamp, and D. Miller (2007): 'Reviewing the literature on interactive whiteboards'. *Learning, Media and Technology*, vol. 32, no. 3, pp. 213–225.

Holmquist, L. E., J. Sanneblad, and L. Gaye (2003): 'Total Recall: In-place Viewing of Captured Whiteboard Annotations'. *NEW HORIZONS*, pp. 2.

Huang, E. M., E. D. Mynatt, D. M. Russell, and A. E. Sue (2006): 'Secrets to success and fatal flaws: the design of large-display groupware'. *IEEE Computer Graphics and Applications*, vol. 26, no. 1, pp. 37–45.

Hummels, C. and J. van Dijk (2015): 'Seven Principles to Design for Embodied Sensemaking'. pp. 21–28, ACM Press.

Inie, N. and P. Dalsgaard (2017): 'How Interaction Designers use Tools to Capture, Manage, and Collaborate on Ideas'. In: *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17*. Denver, Colorado, USA, pp. 2668–2675, ACM Press.

Jaasma, P., J. v. Dijk, J. Frens, and C. Hummels (2017): 'On the Role of External Representations in Designing for Participatory Sensemaking'. In: M. B. Alonso and E. Ozcan (eds.): *Proceedings of the Conference on Design and Semantics of Form and Movement - Sense and Sensitivity, DeSForM 2017*. InTech.

Ju, W. G., B. A. Lee, and S. R. Klemmer (2007): 'Range: exploring proxemics in collaborative whiteboard interaction'. In: *CHI '07 extended abstracts on Human factors in computing systems - CHI '07*. San Jose, CA, USA, p. 2483, ACM Press.

Keller, A. I. (2005): 'For inspiration only: designer interaction with informal collections of visual material.'. Ph.D. thesis, [s.n.], S.l. OCLC: 839640315.

Klemmer, S. R., B. Hartmann, and L. Takayama (2006): 'How bodies matter: five themes for interaction design'. In: *Proceedings of the 6th ACM conference on Designing Interactive systems - DIS '06*. University Park, PA, USA, p. 140, ACM Press.

Klemmer, S. R., M. W. Newman, R. Farrell, M. Bilezikjian, and J. A. Landay (2001): 'The Designers' Outpost: A Tangible Interface for Collaborative Web Site'. In: *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology*. New York, NY, USA, pp. 1–10, ACM.

Klinkhammer, D., M. Mateescu, C. Zahn, and H. Reiterer (2018): 'Mine, Yours, Ours: Coordination Through Workspace Arrangements and Territoriality in Tabletop Interaction'. In: *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia*. New York, NY, USA, pp. 171–182, ACM.

Klokmose, C. N. and O. W. Bertelsen (2013): 'The Mysterious Whiteboard'. In: D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, P. KotzÃľ, G. Marsden, G. Lindgaard, J. Wesson, and M. Winckler (eds.): *Human-Computer Interaction âĂŞ INTERACT 2013*, Vol. 8118. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 37–54.

Korpelainen, E. and M. Kira (2013): 'Systems approach for analysing problems in IT system adoption at work'. *Behaviour & Information Technology*, vol. 32, no. 3, pp. 247–262.

Kruger, R., S. Carpendale, S. D. Scott, and S. Greenberg (2004): 'Roles of Orientation in Tabletop Collaboration: Comprehension, Coordination and Communication'. *Computer Supported Cooperative Work (CSCW)*, vol. 13, no. 5-6, pp. 501–537.

Landay, J. A. and B. A. Myers (1995): 'Interactive Sketching for the Early Stages of User Interface Design', pp. 8.

Larkin, J. H. and H. A. Simon (1987): 'Why a Diagram is (Sometimes) Worth Ten Thousand Words'. *Cognitive Science*, vol. 11, no. 1, pp. 65–100.

Mangano, N., T. D. LaToza, M. Petre, and A. v. d. Hoek (2015): 'How Software Designers Interact with Sketches at the Whiteboard'. *IEEE Transactions on Software Engineering*, vol. 41, no. 2, pp. 135–156.

Mangano, N., T. D. LaToza, M. Petre, and A. van der Hoek (2014): 'Supporting Informal Design with Interactive Whiteboards'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 331–340, ACM.

Maquil, V. and E. Ras (2012): 'Collaborative Problem Solving with Objects: Physical Aspects of a Tangible Tabletop in Technology-based Assessment'. In: J. Dugdale, C. Masclet, M. A. Grasso, J.-F. Boujut, and P. Hassanaly (eds.): *From Research to Practice in the Design of Cooperative Systems: Results and Open Challenges*. London: Springer London, pp. 153–166.

Marshall, P., E. Hornecker, R. Morris, N. S. Dalton, and Y. Rogers (2008): 'When the fingers do the talking: A study of group participation with varying constraints to a tabletop interface'. In: *2008 3rd IEEE International Workshop on Horizontal Interactive Human Computer Systems*. pp. 33–40.

Morris, M. R., J. Lombardo, and D. Wigdor (2010): 'WeSearch: Supporting Collaborative Search and Sensemaking on a Tabletop Display'. In: *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*. New York, NY, USA, pp. 401–410, ACM.

Müller-Tomfelde, C. (ed.) (2010): *Tabletops: horizontal interactive displays*, Human-computer interaction series. London ; New York: Springer. OCLC: ocn495598090.

Mynatt, E. D., T. Igarashi, W. K. Edwards, and A. LaMarca (1999): 'Flatland: New Dimensions in Office Whiteboards'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 346–353, ACM.

Pinelle, D., M. Barjawi, M. Nacenta, and R. Mandryk (2009): 'An Evaluation of Coordination Techniques for Protecting Objects and Territories in Tabletop Groupware'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 2129–2138, ACM.

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. 447–473.

Price, E., S. Tsui, A. Hart, and L. Saucedo (2011): 'Don't Erase that Whiteboard! Archiving Student Work on a Photo-Sharing Website'. *The Physics Teacher*, vol. 49, no. 7, pp. 426–428.

Rebecca, R., G. Joanne, W. Jeremy, G. Peter, P. Alan, H. Stephanie, and D. Dawn (2015): 'Electronic Whiteboards: Review of the Literature'. *Studies in Health Technology and Informatics*, pp. 389–393.

Rick, J., P. Marshall, and N. Yuill (2011): 'Beyond one-size-fits-all: how interactive tabletops support collaborative learning'. In: *Proceedings of the 10th International Conference on Interaction Design and Children - IDC '11*. Ann Arbor, Michigan, pp. 109–117, ACM Press.

Rogers, Y. and S. Lindley (2004): 'Collaborating around vertical and horizontal large interactive displays: which way is best?'. *Interacting with Computers*, vol. 16, no. 6, pp. 1133–1152.

Saldana, J. (2009): 'The Coding Manual for Qualitative Researchers - An Introduction to Codes and Coding'.

Sanders, E.-N. (2000): 'Generative tools for co-designing'. In: *Collaborative design*. Springer, pp. 3–12.

Saund, E. (1999): 'Bringing the Marks on a Whiteboard to Electronic Life'. In: N. A. Streitz, J. Siegel, V. Hartkopf, and S. Konomi (eds.): *Cooperative Buildings. Integrating Information, Organizations, and Architecture*. pp. 69–78, Springer Berlin Heidelberg.

Scott, S. D., M. S. T. Carpendale, and K. M. Inkpen (2004): 'Territoriality in Collaborative Tabletop Workspaces'. In: *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*. New York, NY, USA, pp. 294–303, ACM.

Smit, D., D. Oogjes, B. G. de Rocha, A. Trotto, Y. Hur, and C. Hummels (2016): 'Ideating in Skills: Developing Tools for Embodied Co-Design'. pp. 78–85, ACM Press.

Tang, A., J. Lanir, S. Greenberg, and S. Fels (2009): 'Supporting Transitions in Work: Informing Large Display Application Design by Understanding Whiteboard Use'. In: *Proceedings of the ACM 2009 International Conference on Supporting Group Work*. New York, NY, USA, pp. 149–158, ACM.

Tang, J. C. (1991): 'Findings from observational studies of collaborative work'. *International Journal of Man-Machine Studies*, vol. 34, no. 2, pp. 143–160.

Thatcher, J. B., D. H. McKnight, E. W. Baker, R. E. Arsal, and N. H. Roberts (2011): 'The Role of Trust in Postadoption IT Exploration: An Empirical Examination of Knowledge Management Systems'. *IEEE Transactions on Engineering Management*, vol. 58, no. 1, pp. 56–70.

Thom-Santelli, J., D. Cosley, and G. Gay (2010): 'What Do You Know?: Experts, Novices and Territoriality in Collaborative Systems'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 1685–1694, ACM.

Trier, M. and A. Richter (2013): '" I Can Simply..."-Theorizing Simplicity As A Design Principle And Usage Factor."'. In: *Proceedings of the 21st European Conference on Information Systems*. p. 72.

Triplett, S. M. (2016): 'Design Workspaces: Preserving and Recovering Meaning in Groups of Sketches'. Ph.D. thesis, UC Irvine.

Tuddenham, P. and P. Robinson (2009): 'Territorial coordination and workspace awareness in remote tabletop collaboration'. In: *Proceedings of CHI 2009: ACM Conference on Human Factors in Computing Systems, ACM*. pp. 2139–2148.

van Dijk, J. and G. W. Vos (2011): 'Traces in creative spaces'. In: *Proceedings of the 8th ACM conference on Creativity and cognition*. pp. 91–94, ACM.

Van House, N. A. (2006): 'Interview viz: visualization-assisted photo elicitation'. In: *CHI '06 extended abstracts on Human factors in computing systems - CHI EA '06*. Montreal, Quebec, Canada, p. 1463, ACM Press.

Varona-Marin, D., J. A. Oberholzer, E. Tse, and S. D. Scott (2018): 'Post-meeting Curation of Whiteboard Content Captured with Mobile Devices'. In: *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*. New York, NY, USA, pp. 43–54, ACM.

Wallace, J. R., S. D. Scott, and C. G. MacGregor (2013): 'Collaborative Sensemaking on a Digital Tabletop and Personal Tablets: Prioritization, Comparisons, and Tableaux'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 3345–3354, ACM.

Walny, J., J. Haber, M. Dörk, J. Sillito, and S. Carpendale (2011): 'Follow that sketch: Lifecycles of diagrams and sketches in software development'. In: *2011 6th International Workshop on Visualizing Software for Understanding and Analysis (VISSOFT)*. pp. 1–8.

Zagermann, J., U. Pfeil, R. Rädle, H.-C. Jetter, C. Klokmose, and H. Reiterer (2016): 'When Tablets Meet Tabletops: The Effect of Tabletop Size on Around-the-Table Collaboration with Personal Tablets'. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 5470–5481, ACM.

Zhang, Y., C. J. Yang, S. E. Hudson, C. Harrison, and A. Sample (2018): 'Wall++: Room-Scale Interactive and Context-Aware Sensing'. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. Montreal QC, Canada, pp. 1–15, ACM Press.

Zhang, Z. and L.-w. He (2004): 'Note-taking with a camera: whiteboard scanning and image enhancement'. In: *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 3. Montreal, Que., Canada, pp. iii–533–6, IEEE.

Burrows, R.; Mendoza, A.; Sterling, S.; Miller, T.; Pedell, S. (2019): Evaluating Ask Izzy: A Mobile Web App for People Experiencing Homelessness. In: Proceedings of the 17th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing and the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep17

Evaluating *Ask Izzy*: A Mobile Web App for People Experiencing Homelessness

Rachel Burrows[1,3], Antonette Mendoza[1], Leon Sterling[1,2], Tim Miller[1], Sonja Pedell[2]
[1] School of Computing and Information Systems, The University of Melbourne, [2] Centre for Design Innovation, Swinburne University of Technology and [3] PsyLab, Cambridge Science Park, UK
rburrows@psylab.co.uk,{mendozaa, tmiller}@unimelb.edu.au, {lsterling, spedell}@swin.edu.au

Abstract. This paper contributes to an ongoing discussion in the research community regarding the role of new technology in the lives of those experiencing homelessness. *Ask Izzy* is a mobile web app designed to help people who are homeless or at risk of becoming homeless find the services they need. Since deployment in 2016, it is now attracting over 10,000 users each month. We explore the perceptions towards the design and use of Ask Izzy with a specific focus on emotional concerns. We interviewed 30 participants who were either homeless, ex-homeless, service providers or software developers of the web application. Seven themes emerged from the analysis that appeared to act as barriers or enablers to the uptake of the technology. We discuss how these themes are associated with aspects of technology design or an associated experience with a service provider. We also contrast the views of those who are homeless with service providers. We believe these themes will provoke discussion and be useful for others who are designing for those who are homeless.

Introduction

There are numerous current societal problems that require us to change the way we collectively work together. Well publicised examples include climate change,

population health and wellbeing. These (wicked) problems require solutions that consider the broader socio-technical system in order to address the needs of technology users. One particularly challenging problem is homelessness. In Australia, the number of people experiencing homelessness is up 14% in the five years leading up to 2016 according to the Australian Bureau of Statistics. People without an adequate place to live are likely to be frequently seeking help with service providers for a number of years (Humphry (2014)).

This is a unique, urgent and poorly understood challenge with potential for many technological solutions. As a consequence, a growing body of work in CSCW and HCI is calling for an in-depth understanding of the needs of people who are homeless or at risk of becoming homeless. In previous related work, the design needs of vulnerable user groups has been shown to be a unique situation (Vines et al. (2013)), with some research focusing specifically on homelessness (e.g. Woelfer and Hendry (2011); Hersberger (2013); Chatman (1996); Le Dantec and Edwards (2008); Muñoz et al. (2004); Griffiths and Scarantino (2005)). This work raises many questions. For instance, work by Woelfer and Hendry (2011) questions whether this group of users have the means and motivation to access information online. Work by Hersberger (2013) questions whether this group of users already suffer from information overload from existing service providers and whether new information online will just add to the confusion. Finally, work by Chatman (1996) questions whether the lack of economic independence restricts access to computers and internet resources, and limited access to training hinders uptake of digital technology.

It is clear from prior work that desiging technology to help those who are homeless needs to be approached in a sensitive manner. In particular, information about emotions as discussed by Norman (2013) is important to consider as people often reject technology if it does not support the way they wish to feel while interacting with it. An increasing body of work now focuses on the way a user wants to feel while interacting with technology (e.g. Hou et al. (2017); Pedell et al. (2017); Toscos et al. (2013)). Users may wish to feel in control, connected, hopeful, cared for, or empowered, among others (Toscos et al. (2013); Pedell et al. (2014); Saffarizadeh et al. (2017)). Information about emotions is still extremely difficult to incorporate and evaluate in technology design as they are subjective and situation-dependent. Also, emotional views about technology are formed and change over time based not only on the actual technology engagement experience but also are layered with associated experiences (Saffarizadeh et al. (2017); Alatawi et al. (2018)). Due to the sensitive nature of this application domain, we use information about emotional experiences to guide our analysis and understanding.

Our research is based on a currently deployed mobile web app – called Ask Izzy – that helps homeless Australians find information about the services they need. Specifically, Ask Izzy was launched in 2016. Ask Izzy contains information about services providers, and currently attracts over 10,000 users each month. There are 16 service categories including food, housing, everyday needs, money help and

counselling among others. We conduct semi-structured interviews with users of Ask Izzy. Participants are either homeless, ex-homeless, or service providers. Based on the interview transcripts, we conduct a thematic analysis. We aim to understand the challenges and opportunities that should be considered when designing similar technology. Consequently, the research question we aim to answer in this paper is: *'What are the needs of those experiencing homelessness that should be considered in the design of new technology?'*. The analysis contributes to the existing literature as it provides an in-depth understanding of the technology needs for a unique and poorly understood user group: those experiencing *homelessness*. These results may also contribute to literature on value sensitive design (Friedman et al. (2008)).

1 Background

A common misconception is that technology to help those experiencing homelessness is only for those sleeping rough. The Australian Bureau of Statistics defines a person experiencing homelessness as

> "...in a dwelling that is inadequate; or has no tenure, or if their initial tenure is short and not extendable; or does not allow them to have control of, and access to space for social relations."

This typically means that people who are in shelters or transitional accommodation that has been purposefully built for homeless people are also considered to be homeless. The goal of reaching a stable housing situation may take a number of years, and involve frequent interactions with multiple service providers. In this section we discuss related work on technology for homelessness and the importance of considering emotions in technology design.

1.1 Technology for Homelessness

There are many considerations when designing technology for those experiencing homelessness. Currently, most information is exchanged in face-to-face situations (Hersberger (2013); Le Dantec and Edwards (2008)). Providing information online about available services is one way to increase accessibility and help those that are homeless to search and find the help that they need. Prior research has questioned whether access to additional information online is effective. Work by Hersberger (2013) questioned whether those who are homeless are not already overwhelmed by information provided by services. Additionally, the lack of economic independence restricts access to computers and internet resources. Equally, limited access to training hinders uptake of digital technology (Chatman (1996)). One study by Woelfer and Hendry (2011) suggests that we should take a precautionary stance when it comes to providing access to complex service information online and has even suggested that ubiquitous technology may not be the solution to this problem. Additionally, service providers often resist change and reduction in

control over how those who are homeless are accessing information. This is because the new means of accessing information creates expectations upon service providers that they may not be able to meet (Le Dantec and Edwards (2008); Weise et al. (2017)).

Despite the scepticism regarding the effectiveness of communicating complex service information online, there is evidence to suggest it would be beneficial to many. A mobile phone is often viewed as a necessity and a critical lifeline (Le Dantec and Edwards (2008)). A recent study estimated most people experiencing homelessness prioritise retaining their mobile phone to keep in touch with family, friends and necessary service providers (Humphry (2014)). Even those without a smart phone may have access to the internet via alternative means, such as the library or with help from case workers. More recent work by Woelfer and Hendry (2012) investigates the extensive use of social media by young homeless people, highlighting different information seeking strategies utilised by a younger tech-savvy generation.

Those experiencing homelessness represent a unique user group. The major causes of homelessness are outlined in Table I and include family violence, financial difficulties or a housing crisis. Living situations are outlined in Table II. For those in these situations, additional problems often accumulate over time, such as drug and alcohol abuse, creating a viscous cycle (Woelfer and Hendry (2009)). In reaction to these complex needs, services allocate an extensive range of support, and become largely responsible for the diffusion of new information and support to those who are homeless via a mix of government funded organisations and grassroots organisations (Woelfer and Hendry (2009)).

Table I: Causes

Reason	%
Family violence	24
Financial difficulties	20
Housing crisis	16
Inadequate dwellings	11
Other relationship issues	8
Other housing issues	5
Health issues	4
Other	12
Reference for table data:	
Homelessness Australia;	
ABS;	
Chamberlain et al. (2014)	

Table II: Living Situations

Place	%
Severe overcrowding	39
Supported accommodation	20
Temporarily staying with others	17
Boarding houses	17
Improvised / rough sleeping	6
Other temporary lodging	1

1.2 Considering Emotions in Technology

People will reject new technology if it does not appeal to their emotional needs (Dix et al. (2003); Krumbholz et al. (2000); Norman (2005); Pedell et al. (2014); Miller et al. (2015)). For this reason, there is a growing body of work that aims to use information about emotions of users to improve the design of new technology. In our study, we are not directly measuring an emotional state. Instead we are using information about emotional experiences discussed in interviews as a basis for evaluating the design of a mobile web app for homelessness. We discuss this further in the following section.

While there are a variety of popular psychological frameworks that characterise emotions, their content and utility for system design and evaluation will vary. In this section we give an overview of popular psychological frameworks that do characterise emotion and its influence on technology use.

Some psychological frameworks are grounded in primary (also referred to as basic) emotions such as fear, anger, or joy (Ekman (1992); Schwarz and Clore (1983)). These frameworks can then be used by technology developers to evaluate whether such emotions are incorporated into the technology itself (Sutcliffe (2009); Lowry et al. (2012)). Other emotional frameworks contain different types of emotions, including those that are more reflective. For instance, some emotions are characterised by having relatively lower levels of arousal and involve relatively higher levels of reflective, cognitive processes; examples include the characterisation of shame and resentment (Martin and Tesser (1996); Desmet and Hekkert (2007); Plutchik (2003)).

Emotional experiences associated with technology use may be related to aspects of the software design, such as a particular feature that is displayed. Work on socio-materiality and technology affordances (Orlikowski and Scott (2008); Majchrzak et al. (2013); Vaast and Kaganer (2013); Leonardi (2013)) shows how aspects of design can trigger positive and negative emotional perceptions. Emotional experiences associated with a particular technology are also influenced by external factors, including other individuals or organisations that are associated with the engagement experience. Misplaced expectations may still be attributed to the technology itself due to multiple experiences becoming aggregated and associated with each other (Wood and Moreau (2006)). These experiences could include those occurring during the progression towards a common goal (Clore and Ortony (2008); Luce et al. (2001)) or achievement (Martin and Tesser (1996)). Consequently, in our case study, participants may attribute perceptions related to an interaction with a service provider wrongly to the technology that facilitated the interactions.

Those that are homeless experience a range of different emotions that place poorly understood demands upon creators of new technology. Unfortunately, while there are some studies that focus on designing for vulnerable user groups, and even those who are homeless, these studies do not focus specifically on emotions and also do not evaluate a deployed system that has been designed with these needs in mind.

2 Case Study

2.1 A Mobile Web App for Homelessness

Ask Izzy (Infoxchange (2018b)) is a mobile web app that aims to tackle the problem of homelessness by assisting those who are homeless in finding useful information. It provides information about services in Australia. The listed services provide a range of support from help with health issues, food, shelter through to legal and financial advice. The website was listed as 'un-metered' with the network; meaning, it does not cost anything to access. Additionally, battery packs were donated to increase the capability of those in need to use their phones for longer.

A typical use of Ask Izzy involves starting at the landing page shown in Figure 1. A user is presented with 16 *help categories*. Examples of the categories are Housing, Food and Money Help. The user can choose to give their location and is guided through a series of *category-specific questions*. Based on these answers, the user is presented with a *service list* page compiled via a *service filter process* detailing results of services that match their criteria, and ordered by relevance. A user can select a particular service and view its *detailed service page*. The *detailed service page* (Figure 1) displays information about how to connect with the particular service, how to get there, who it is for and what clients should expect.

2.2 Method

Two authors conducted a series of semi-structured, one-hour interviews with 30 participants six months after the deployment of Ask Izzy. We took care to ensure the recruitment procedures and interview locations were appropriate; participants were recruited via existing service providers who were also able to provide a familiar environment for the discussion to take place.

Participants who were homeless had some experience with Ask Izzy. This ranged from a single use to frequent use over the 6 months period of time. Service providers were aware of Ask Izzy and therefore played a role in raising awareness with homeless people with whom they were in contact. Table III gives an overview of the participants that were interviewed. Participants were selected to represent a range of people who have a stake in the success of the application and who have had first hand experience with Ask Izzy. This included those who were homeless, ex-homeless, service providers, and the software company, in a range of situations.

A semi-structured interview was chosen to give flexibility to the conversation. It allowed participants to diverge and discuss contextual factors that may be unexpectedly related to their perceptions of Ask Izzy. With regard to the software design, we asked what they liked, disliked, and what they would change in the mobile app. We also asked how using Ask Izzy made them feel. We discussed interactions and experiences outside of the application including how they heard about Ask Izzy and if they had recommended it or supported others in using it. We also asked what they thought were the barriers to uptake. If they chose not to use Ask Izzy, we asked for the reason.

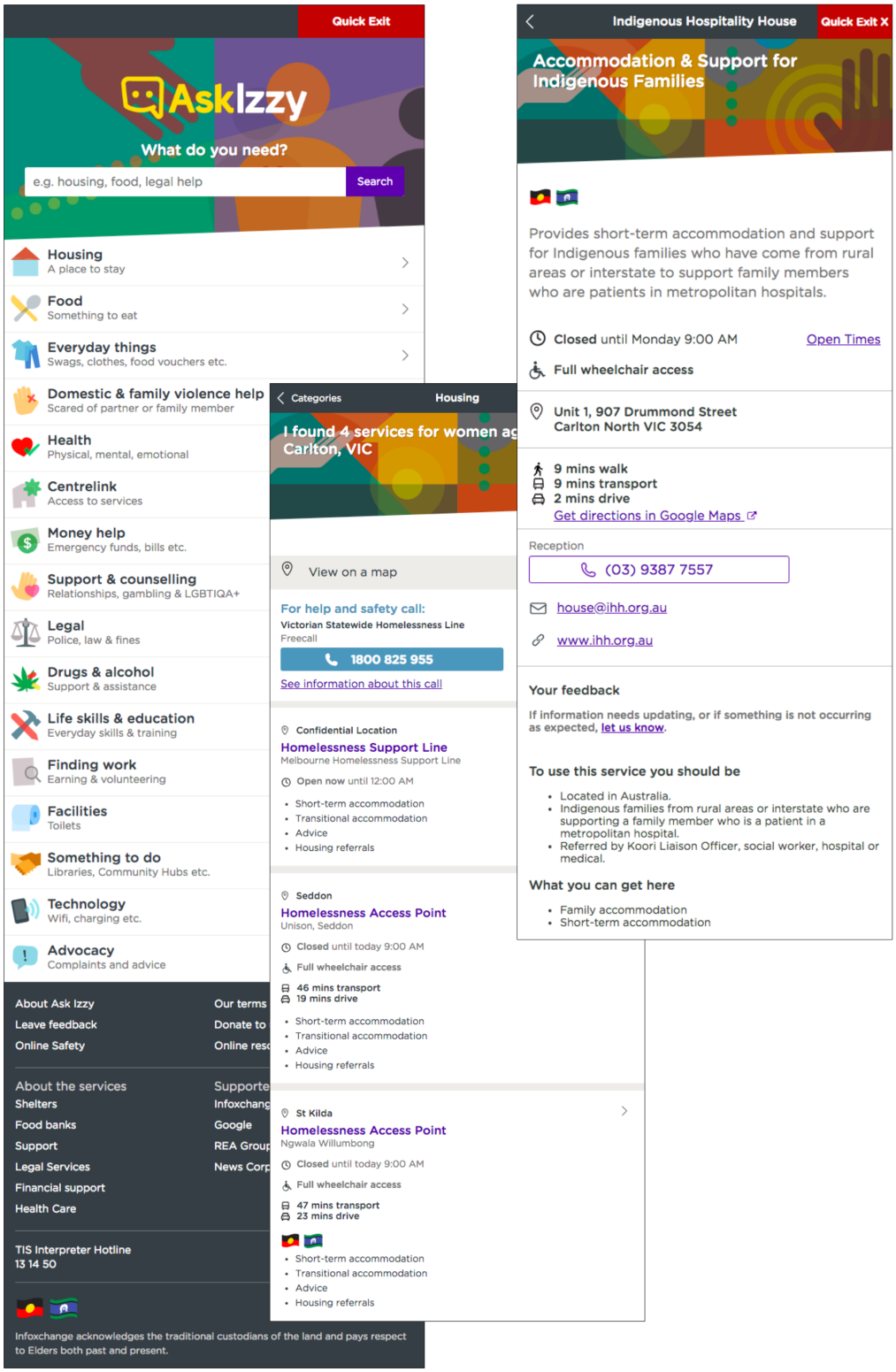


Figure 1: A sample of screenshots from the second release of Ask Izzy. Left to right: Landing page, service list page and detailed service page.

Table III: Interview Participants

Participant Groups	Number	Sample Coverage
Homeless and Ex-Homeless	14	Adult, Youth, Family Violence, Veteran, Mental or Emotional Difficulties, Drugs and Alcohol Problems, With Children Stable Living Conditions, Unstable Living Conditions
Service Providers	15	Official Service Providers including Government funded and Charities, inc. Services Providers for Aboriginal and Torres Strait Islanders
Software Owner	1	A representative from the software company

All transcripts and audio recordings were imported into the NVivo tool (Bazeley and Jackson (2013)). The results were analysed by two authors following a thematic data analysis process (Braun (2006)) in order to identify, analyse, and report the themes from the data. Codes (i.e. quotes) were extracted from the transcripts that were related to emotional experiences. The codes of the transcripts were grouped individually into themes and then later discussed and merged to form a final agreed set of themes.

Following guidelines for thematic analysis, no specific framework of emotions was used to categorise the elicited codes. Any text phrases that were thought of as representing the way the participant would or would not like to feel were marked and extracted. We retained information about the role of each participant in order to contrast views of service providers with those who are homeless or ex-homeless.

3 Results

The reaction to Ask Izzy was positive. Seven themes emerged that represented aspects of design that may act as barriers or enablers to the uptake and use of Ask Izzy.

More specifically, 107 codes were extracted from the transcripts and subsequently grouped into the 7 themes. These themes were Empowerment and Control, Assurance, Cared For, Identity and Belonging, Clarity, Unashamed / Without Stigma and Hopeful. We give a brief description of these themes in this section. Note that those who are homeless or at risk of being homeless are referred to as clients by the service providers.

Empowerment and Control: *The empowerment and control theme is about new ways clients can access service information.*
23 codes were extracted and categorised as being related to empowerment and feeling in control. This theme emerged from discussions about the visibility and ordering of services presented on the *service list* page. Ask Izzy provided new ways of accessing services that would not have previously been publicly available.

Assurance: *The theme of being assured is associated with the ways in which trusted information is accessed.*
16 quotes were categorised as being related to feeling assured in the information accessed through the application. This theme was associated with the person or organisation that provides information and the impact that has on confidence in that information.

Cared for: *The theme of feeling cared for was associated with both software features and related social interactions*.
Five codes were categorised and were related to clients feeling cared for when they were provided with useful information. They also discussed how they used Ask Izzy to care for others by accessing it on their behalf.

Identity and Belonging: *Signalling that services are inclusive gives rise to a sense of belonging.*
12 codes were extracted and categorised as being related to a sense of belonging and a sense of identity. This theme was mostly relevant to Aboriginal and Torres Strait Islander user group as they wished to feel that this software application had been designed with their needs in mind.

Clarity: *Clarity is a theme about presenting relevant information without being overwhelming.*
16 codes were extracted and categorised as being related to having clarity as opposed to overwhelming the user who may be stressed or emotionally unstable at that point in time. This theme was associated with a variety of sources relating to understanding the available service options.

Being Unashamed: *Being unashamed is a theme about avoiding the stigma that hinders help-seeking behaviour.*
13 codes were extracted and categorised as being related to the stigma associated with homelessness. Due to the stigma associated with being homeless, clients are apprehensive about asking for help. This theme emerged from discussions around alternative and anonymous means of accessing information.

Hopefulness: *Hopefulness is a theme associated with help-seeking behaviour while managing expectations.*
22 codes were categorised in the theme of hopefulness. Many participants

emphasised the balance in technology design about motivating user to find help while simultaneously managing user expectations.

4 Findings and Discussion

We now revisit our initial research question: *What are the needs of those experiencing homelessness that should be considered in the design of new technology?*

In the following we discuss the seven themes and contrast views of those who are homeless or ex-homeless with those who are service providers. These themes represent design considerations that were found to be positively addressed in the design. However, these themes still had the potential to become barriers to the uptake of Ask Izzy for a variety of reasons.

4.1 Empowerment and Control

The extent to which service providers and clients felt in control was influenced by the number and types of services that were listed in the *service list* page. Concretely, people who were homeless explained how the mobile web app frequently presented more service options than they were previously aware of and that the power to choose which one to access was in their hands.

> "*... you don't have to go to that one, you can have a choice.*"

Contrary to the views of clients, the service providers had a different stance. Ask Izzy would reduce the control that service providers had regarding the ways in which clients access the services. They were concerned that clients would accidentally be provided with inaccurate information while searching for services, and consequently end up approaching an organisation that was not able to help. One service provider stated:

> "*It's worse knowing it's there and that they're not going to be able to help me. It would be better thinking there is only one service.*"

This tension mirrors findings in related work (e.g. Weise et al. (2017)) that has documented the changing shift in power from governments and service providers to the public. What underlay this tension was the ability to access service providers via searching for their details online as opposed to a recommendation by an existing case worker or service provider. Consequently, the information that is presented in Ask Izzy came with a risk that a homeless person would attempt to access a service that was inappropriate to their situation. For example, one participant stated that information about service providers is sometimes only provided by a referral via another service provider. This prevents many people from even knowing that the service exists but gives the service provision network the control to only recommend this service to those that would qualify to receive it.

4.2 Assurance

Those who are experiencing homelessness trust recommendations from others in the same situation, and as such, trust is a vital mechanism of discovering new information. One person experiencing homelessness said:

> "*Word of mouth, word of mouth, whatever they hear on the streets. So they take their opinions and advice of people, other people that are homeless that have been there and know the system. Who's who. [...] Because people don't let you down, on that side of life.*"

This illustrates the power of trusted information sharing within a community of people experiencing homelessness. The trust in whoever is recommending Ask Izzy, and also past experiences with services, are therefore transferred to trust in the application itself.

4.3 Cared for

While participants were reflecting on their interaction with Ask Izzy they described how the language that guided them to find their service was personable. Ask Izzy was created for the purpose of helping them and they consequently felt cared for. However, the extent to which a client felt cared for was dependent on the accumulated interactions with service providers, those others who recommend or who are accessible via Ask Izzy.

Another client explained how he frequently used Ask Izzy to help others:

> "*There's a lot that have come up to me and go 'we haven't got smartphones but you've got your Ask Izzy' [...] A lot of them come back and go oh that was very positive, where else can we go to?'*."

This is an example where the subsequent interaction with the newly found service provider influences how much a particular client feels cared for. For this reason, the feeling of being *cared for* may change with each experience seeking help. In short, the design creates the expectation of feeling cared for and is therefore strongly influenced by the series of interactions with multiple service providers over what is likely to be numerous years.

This example also illustrates an interaction with Ask Izzy is not necessarily one person with a mobile device who is helping (caring for) another to find what they need. This interaction was not the primary way that designers envisaged Ask Izzy to be used, however, our findings indicate that many interactions with Ask Izzy are social. They may be between two people that are currently homeless where one is an expert user. Other engagements involved a service provider finding details on behalf of a client.

4.4 Identity and Belonging

This theme was mostly relevant to Aboriginal and Torres Strait Islander user group as they wished to feel that this software application had been designed with their needs in mind. This was a challenge as the interface design preferences from different cohorts of users were very different. The Aboriginal and Torres Strait Islander communities wanted the design to show a signal that their needs had been considered, and to connect them culturally to their community.

They initially described Ask Izzy as "*too mainstream, too governmental, whichever way you want to put it.*". This negative perception was generally related to the aspect of the design that was "*lacking*" as opposed to one that already existed. The feedback became a high priority requirement for future iterations of design to see how the need of identity and belonging could be better addressed.

One representative of the software development team said:

> "*How do you take something that a lot of people like at the moment and then come up with a next generation and then make sure that the things people like are still there [...] That's a really interesting design challenge.*"

4.5 Clarity

Ask Izzy was deliberately designed to communicate the thousands of services as clearly as possible. One participant who had experienced homelessness said:

> "*Your emotions are high and all that sort of stuff you're going through with something you've never experienced in your life before. So from that point of view it's absolutely brilliant, 'cause it tells you, you open it up, and it literally tells you which tram to get onto and which stop to get off and which train and all that sort of and so on.*"

The simplicity of the categories, imagery, icons and language was designed with clarity in mind. When in stressful situations, some clients rely on recommendations from service providers or their case worker to be able to use the application, "*I just talk to my case worker because they have all this information,...*". In these scenarios, Ask Izzy would potentially be used by both clients and service providers together. One service provider explained how they had a link to Ask Izzy on their desktop.

A second sub-theme related to clarity is about the clarity of the purpose of who Ask Izzy was for. Many target users of Ask Izzy do not identify with being homeless. Rough sleeping is heavily stigmatised. Those that have stabilised their living situations are quick to reject the characterisation of their situation of homeless.

One service provider stated:

> "*So that's part of the problem with promoting things as homelessness is most people don't identify.*"

4.6 Being Unashamed

Due to the stigma attached with asking for help, many people who were homeless preferred to access information anonymously through sites like Ask Izzy, or alternatively by creating a fake profile on social media.

One client stated:

> "*Well I find Facebook easier [than speaking in person] because I can be anonymous on Facebook so like I can make up a fake profile and just ask random questions on a group and they can like reply, so it's sort of like word of mouth but it's like word of mouth I don't have to [be there in person]* "

The increasing desire to seek help online was also documented in prior work by Woelfer and Hendry (2012). While there is existing evidence to show that information about services is complex and overwhelming, there is also a growing body of work to show that this means of accessing information is convenient and in the example above, preferred.

4.7 Hopefulness

The use of Ask Izzy becomes a trigger where the hope is created, it then may be acted upon by approaching a service. The amount of hope that a user may feel changes with each subsequent interaction with the app and also service providers. The reality is that the journey to a stable living situation may be a number of years. The greater the initial hope, the greater the risk of negative consequences in the long-term when expectations are potentially not met.

In our interviews, service providers explained how the application needs to set realistic expectations:

> *"It's not a silver bullet in that sense, so I think that, it connects people to information quickly which is really good. But it doesn't necessarily resolve what they need.*"

Feelings of hope on part of the clients were increased with any indication that they may have found a solution to their problems. Too much hope comes with the danger of disappointment when the actual service fails to meet a user's expectations in reality. Too little hope may be caused by explaining the harsh reality of what to expect from under-funded services, and therefore may discourage clients from seeking and/or engaging in social networking services in the first place. Framing the right message to manage client expectations is a challenge here and can sway clients to take action and access services or not. Many clients could recall an instance where they were devastated after a particular service could not help them.

4.8 Limitations

There are some limitations that we took steps to mitigate. Firstly, the thematic analysis comes with the inherent characteristic of those themes being open to a subjective interpretation by coders. To reduce the risk we followed recommended steps of comparing and merging results with more than one coder. Secondly, the results are only for one case study and therefore they may not be representative of other new technology for homelessness. We are therefore cautious with the generalisability of these results as the capability to appropriate technology similar to Ask Izzy is likely to be dependent on multiple societal factors, such as familiarity with and appropriation of e-government services. Despite this, the investigation was grounded in a large industrial case study that aimed to help those who are homeless with a diverse range of everyday needs.

5 Conclusion

People who are homeless are now increasingly tech-savvy and use the internet to find information and ask questions anonymously. We evaluate an existing mobile web app to better understand the technology needs of those experiencing homelessness. We firstly conduct semi-structured interviews with 30 people who were either homeless, ex-homeless, a service provider or a software company. Seven themes emerged from the analysis that would be useful considerations for the design of technology for those experiencing homelessness. These were: empowerment and control, hopefulness, assurance, cared for, identify and belonging, clarity and being unashamed.

During the thematic analysis, we found our focus on emotion useful to gain a shortcut to important contextual information uncovering barriers or enablers to the use of similar technology. In some cases our results confirm and elaborate on design challenges that have been articulated elsewhere in related work. For instance, access to new information via Ask Izzy is empowering for those who are homeless. Also, the need for clarity is important, especially with those who are younger as they can be easily overwhelmed by the amount of information available and the complexities of navigating the service provision network. Our evaluation also offers multiple new insights; many uses of Ask Izzy were *social*. Users with smartphones frequently accessed Ask Izzy on behalf of others who were homeless without a smartphone. Ask Izzy was also used by service providers who used it as a reference point on behalf of those asking a question in-person.

Acknowledgments

This research has been made possible by the support and collaboration with Infoxchange a not-for-profit social enterprise developing technology for social change. This project was

funded by the Australian Research Council Discovery Grant DP160104083 Catering for individuals' emotions in technology development.

References

Alatawi, E., A. Mendoza, and T. Miller (2018): 'Psychologically-Driven Requirements Engineering: A Case Study in Depression Care'. In: *2018 25th Australasian Software Engineering Conference (ASWEC)*. pp. 41–50.

Australian Bureau of Statistics, 'Census of Population and Housing: Estimating homelessness, 2016'. http://www.abs.gov.au/ausstats/abs@.nsf/mf/2049.0.

Bazeley, P. and K. Jackson (2013): *Qualitative Data Analysis with NVivo*. Sage Publications Ltd.

Chamberlain, C., G. Johnson, and C. Robinson (2014): *Homelessness in Australia*. UNSW Press.

Chatman, E. A. (1996): 'The impoverished life-world of outsiders'. *Journal of the American Society for information science*, vol. 47, no. 3, pp. 193–206.

Clore, G. L. and A. Ortony (2008): 'Appraisal theories: How cognition shapes affect into emotion.'.

Desmet, P. and P. Hekkert (2007): 'Framework of product experience'. *International journal of design*, vol. 1, no. 1.

Dix, A., J. E. Finlay, G. D. Abowd, and R. Beale (2003): *Human-Computer Interaction (3rd Edition)*. Prentice-Hall, Inc.

Ekman, P. (1992): 'An argument for basic emotions'. *Cognition & emotion*, vol. 6, no. 3-4, pp. 169–200.

Friedman, B., P. H. Kahn, and A. Borning (2008): 'Value sensitive design and information systems'. *The handbook of information and computer ethics*, pp. 69–101.

Griffiths, P. E. and A. Scarantino (2005): 'Emotions in the wild: The situated perspective on emotion'.

Hersberger, J. (2003): 'A qualitative approach to examining information transfer via social networks among homeless populations'. *The New Review of Information Behaviour Research*, vol. 4, no. 1, pp. 95–108.

Hersberger, J. A. (2013): 'Are the economically poor information poor? Does the digital divide affect the homeless and access to information?'. In: *Proceedings of the Annual Conference of CAIS/Actes du congrès annuel de l'ACSI*.

Homelessness Australia Website (2018): 'http://www.homelessnessaustralia.org.au: accessed 15/01/2019'.

Hou, Y., C. Lampe, M. Bulinski, and J. J. Prescott (2017): 'Factors in Fairness and Emotion in Online Case Resolution Systems'. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. pp. 2511–2522.

Humphry, J. (2014): *Homeless and Connected: Mobile phones and the Internet in the lives of homeless Australians*.

Infoxchange (2018a): 'Inclusive Design'. https://www.infoxchange.org/au/news/2017/10/new-inclusive-design-ask-izzy accessed 15/01/2019.

Infoxchange, I. (2018b): 'https://www.infoxchange.org/au/ask-izzy: accessed 15/01/2019'.

Krumbholz, M. a., J. Galliers, N. Coulianos, and N. Maiden (2000): 'Implementing enterprise resource planning packages in different corporate and national cultures'. *Journal of Information Technology*, vol. 15, no. 4, pp. 267–279.

Le Dantec, C. A. and W. K. Edwards (2008): 'Designs on dignity: perceptions of technology among the homeless'. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. pp. 627–636.

Leonardi, P. M. (2013): 'Theoretical foundations for the study of sociomateriality'. *Information and Organization*, vol. 23, no. 2, pp. 59–76.

Lowry, P. B., J. Gaskin, N. Twyman, B. Hammer, and T. Roberts (2012): 'Taking fun and games seriously: Proposing the hedonic-motivation system adoption model (HMSAM)'.

Luce, M. F., J. R. Bettman, and J. W. Payne (2001): 'Emotional decisions: Tradeoff difficulty and coping in consumer choice'. *Monographs of the Journal of Consumer Research*, no. 1, pp. 1–209.

Majchrzak, A., S. Faraj, G. C. Kane, and B. Azad (2013): 'The Contradictory Influence of Social Media Affordances on Online Communal Knowledge Sharing'. *Journal of Computer-Mediated Communication*, vol. 19, no. 1, pp. 38–55.

Martin, L. L. and A. Tesser (1996): 'Some ruminative thoughts'. *Advances in social cognition*, vol. 9, pp. 1–47.

Massimi, M., J. P. Dimond, and C. A. Le Dantec (2012): 'Finding a new normal: the role of technology in life disruptions'. In: *Proceedings of the acm 2012 conference on computer supported cooperative work*. pp. 719–728.

Miller, T., S. Pedell, A. A. Lopez-Lorca, A. Mendoza, L. Sterling, and A. Keirnan (2015): 'Emotion-led modelling for people-oriented requirements engineering: the case study of emergency systems'. *Journal of Systems and Software*, vol. 105, pp. 54–71.

Muñoz, M., C. Vázquez, and J. J. Vázquez (2004): 'A comparison between homeless, domiciled and vulnerable populations in Madrid'. *Population*, vol. 59, no. 1, pp. 129–142.

Norman, D. A. (2005): *Emotional design: Why we love (or hate) everyday things*. Basic books.

Norman, D. A. (2013): *The design of everyday things: Revised and expanded edition*. Basic books.

Orlikowski, W. J. and S. V. Scott (2008): '10 sociomateriality: challenging the separation of technology, work and organization'. *Academy of Management Annals*, vol. 2, no. 1, pp. 433–474.

Pedell, S., A. Keirnan, G. Priday, T. Miller, A. Mendoza, A. Lopez-Lorca, and L. Sterling (2017): 'Methods for Supporting Older Users in Communicating Their Emotions at Different Phases of a Living Lab Project'.

Pedell, S., T. Miller, F. Vetere, L. Sterling, and S. Howard (2014): 'Socially-Oriented Requirements Engineering: Software Engineering Meets Ethnography'. In: *Perspectives on Culture and Agent-based Simulations*. Springer, pp. 191–210.

Plutchik, R. (2003): *Emotions and life: Perspectives from psychology, biology, and evolution.* American Psychological Association.

Saffarizadeh, K., M. Boodraj, and T. M. Alashoor (2017): 'Conversational Assistants: Investigating Privacy Concerns, Trust, and Self-Disclosure'.

Schwarz, N. and G. L. Clore (1983): 'Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states.'. *Journal of personality and social psychology*, vol. 45, no. 3, pp. 513.

Sutcliffe, A. (2009): 'Designing for user engagement: Aesthetic and attractive user interfaces'. *Synthesis lectures on human-centered informatics*, vol. 2, no. 1, pp. 1–55.

Toscos, T., K. Connelly, and Y. Rogers (2013): 'Designing for positive health affect: Decoupling negative emotion and health monitoring technologies'. In: *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2013 7th International Conference on*. pp. 153–160.

Vaast, E. and E. Kaganer (2013): 'Social media affordances and governance in the workplace: An examination of organizational policies'. *Journal of Computer-Mediated Communication*, vol. 19, no. 1, pp. 78–101.

Vines, J., R. McNaney, R. Clarke, S. Lindsay, J. McCarthy, S. Howard, M. Romero, and J. Wallace (2013): 'Designing for-and with-vulnerable people'. In: *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. pp. 3231–3234.

Weise, S., P. Coulton, and M. Chiasson (2017): 'Designing in between Local Government and the Public–Using Institutional Analysis in Interventions on Civic Infrastructures'. *Computer Supported Cooperative Work (CSCW)*, vol. 26, no. 4-6, pp. 927–958.

Woelfer, J. P. and D. G. Hendry (2009): 'Stabilizing homeless young people with information and place'. *Journal of the American Society for Information Science and Technology*, vol. 60, no. 11, pp. 2300–2312.

Woelfer, J. P. and D. G. Hendry (2011): 'Designing ubiquitous information systems for a community of homeless young people: precaution and a way forward'. *Personal and Ubiquitous Computing*, vol. 15, no. 6, pp. 565–573.

Woelfer, J. P. and D. G. Hendry (2012): 'Homeless Young People on Social Network Sites'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 2825–2834, ACM.

Woelfer, J. P., A. Iverson, D. G. Hendry, B. Friedman, and B. T. Gill (2011): 'Improving the safety of homeless young people with mobile phones: values, form and function'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 1707–1716.

Wood, S. L. and C. P. Moreau (2006): 'From fear to loathing? How emotion influences the evaluation and early use of innovations'. *Journal of Marketing*, vol. 70, no. 3, pp. 44–57.

Farshchian, B.A. (2019): Collaboration as Commodity: What does CSCW have to offer? In: Proceedings of the 17th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing an the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep09

Collaboration as Commodity: What does CSCW have to offer?

Babak A. Farshchian
Norwegian University of Science and Technology (NTNU)
Contact Author: Babak.Farshchian@ntnu.no

Abstract. CSCW as a research field has contributed to the development of digital tools and platforms to support collaboration. Historically, detailed studies of collaboration have played a central role in the development of theories in CSCW. Parties to collaboration have been the main actors, engaged in synchronous or asynchronous, co-located or distant collaboration. CSCW has often considered the platform, i.e., the place where data about collaboration is stored, as a neutral actor without own agency or agenda. This picture has however changed drastically with the recent emergence of digital labor platforms and data-driven business models. Digital labor platforms move the focus from collaborating actors to platform owners, from supporting collaboration to trading collaboration as commodity. In this paper, I attempt to describe this development from a CSCW perspective. I propose a way to re-frame existing knowledge to fit into the new paradigm of collaboration as commodity. I propose to use research from neighboring fields such as information systems to increase our impact as CSCW researchers. Finally, I discuss several research questions for CSCW. This is work in progress.

Introduction

My aim in this paper is to raise a discussion about the role of CSCW in the new landscape of *digital labor platforms* for the so-called gig or on-demand economies (Choudary, 2018; Frenken & Schor, 2017). Digital labor platforms – interchangeably called platforms in the rest of this paper –are IT-based online services that create a market for labor and facilitate its trade online. These platforms

allow individual consumers and workers to find each other, and to buy and sell labor. Some of these platforms have become global powerhouses –Uber, Lyft, Amazon Mechanical Turk, UpWork etc. –and connect tens of millions of consumers and workers across continents. Several such platforms support mainly online work –e.g. Amazon Mechanical Turk –while others combine online and physical offline collaboration –e.g. Uber. Increasingly new types of labor are transformed into on-demand models supported by digital labor platforms. Between 9% and 13% of the population in several European countries report being frequent platform workers (Huws, Spencer, & Joyce, 2016).

So-called "platform models" (Tiwana, 2013) have attracted considerable attention among researchers in the fields of information systems and management (Constantinides, Henfridsson, & Parker, 2018; de Reuver, Sørensen, & Basole, 2018). Numerous CSCW researchers have also studied these platforms and their users under terms such as micro-task (Gupta, Martin, Hanrahan, & O'Neill, 2014), crowdsourcing (Gray, Suri, Ali, & Kulkarni, 2016), and on-demand or labor platforms (Harmon & Silberman, 2018). Such studies have shed valuable light on how users perceive and use digital labor platforms.

CSCW research has resulted in debate about current global labor platforms and the way users –in particular workers –are treated by such platforms. CSCW research often addresses the interactions among platform users, while some emerging studies also look at the interaction between platform users and owners, such as (Glöss, McGregor, & Brown, 2016; Harmon & Silberman, 2018; Kittur et al., 2013). Overarching models and theories to study and debate platform ecosystems –in particular, the interactions between platform owners and users –are so far less emphasized in CSCW research. Therefore, design implications that result from mainstream CSCW research often do not question the fundamental governance mechanism and business models inherent in current platforms.

At the same time, CSCW researchers are in an excellent position to impact the design and evolution of labor platforms based on our focus and understanding of how users perceive and use digital labor platforms. Impact based on a knowledge of users needs to be systematic and result in practical advice and design considerations if they are to be used by platform owners. Such an impact does not need to contradict the interests of platform owners. Examples of costly legal and reputational battles that global labor platforms are currently fighting are abundant (Rodes, 2017; Semuels, 2018). A CSCW research agenda taking into consideration the whole ecosystem of digital labor platforms can make itself relevant not only for user representatives –such as labor unions –but also for platform owners who struggle with a poor reputation and associated costs due to poor design choices and governance models.

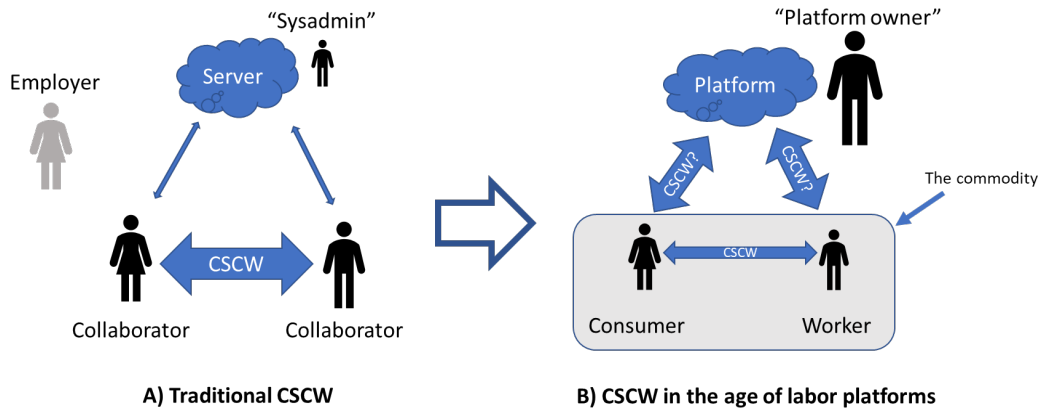


Figure 1: How platforms change the balance from collaborators to platform owners.

Digital labor platforms represent a fundamental transition for CSCW research. This transition is illustrated in Figure 1. Conventional CSCW (shown in the left side of the figure) focuses on interactions among collaborators and has generated a multitude of theories to explain and digitally support such interactions. Data about collaboration has always been central in CSCW –awareness data, context data, coordination data, etc. Historically, the system administrator – "sysadmin" –was a taken for granted agent who administered a "server" where this data was stored and accessed by collaborators and their tools. Sysadmin was often regarded as an actor without any political agenda, and his/her role has seldom been a subject of CSCW research. Employers, who on the other hand might have a political agenda –e.g. maximizing profit –have neither been the focus of mainstream CSCW research but their relationship with employees has been studied by e.g. participatory design researchers (Bratteteig & Wagner, 2016).

Fast-forward to the new landscape of online platforms, and we have a scenario where a digital labor platform has replaced the server, and the employer and the sysadmin have taken on a new join role as the *platform owner* (shown to the right side of Figure 1). In this new scenario, the collaboration between collaborators is transformed into a transaction –or a series of transactions –between a *worker* and a *consumer* of labor. Seen from the platform owner's perspective, collaboration in the conventional CSCW sense –and its outcome –constitutes a *commodity* that can be traded. To facilitate its trade, collaboration needs to be simplified and standardized. The details of the interactions between consumers and workers –which have traditionally been the core area for CSCW researchers –are of interest to the platform owner as far as these details can contribute to generating revenues for the platform. Platform owners –through their governance models as we will discuss later –restrict and guide these interactions with the aim of increasing platform revenues.

My argument in this paper is that CSCW researchers need to pay closer attention to this transition from "server" to "platform," and the complex ecosystem that has

emerged among platform owners, workers and consumers. So far, most CSCW research on labor platforms, micro-task platforms, crowdsourcing, etc., has been concerned with how collaboration is done among workers and consumers, e.g. (Glöss et al., 2016; Kittur et al., 2013; Raval & Dourish, 2016). Additionally, CSCW researchers have recently created a research agenda promoting a quantitative view of collaboration. We see emerging studies that abstract away from the collaboration itself and look at its macro aspects. For instance, Hata et al. (Hata, Krishna, Li, & Bernstein, 2017) investigated long term worker fatigue and its effect on the quality of results among large groups of Amazon Mechanical Turk workers. Ahmed and Fuge (Ahmed & Fuge, 2017) used algorithms to discover and select high-quality ideas from mass online collaboration. De Boer and Bernstein (de Boer & Bernstein, 2017) used statistical models to identify well-performing crowd processes given a business objective. I believe this strand of research fails to build on the strength of CSCW in studying details of work practices. On the other hand, there is some emerging research that questions the relationship and the (lack of) collaboration between platform owners and platform users (Gupta et al., 2014). A coherent research agenda can increase CSCW's impact on how digital labor platforms are developed and evolve.

In the rest of this paper, I will first introduce some background on collaboration as commodity, and two concepts from information systems literature, i.e., platform governance models and boundary resources, that in my view can help structure existing research in a new light. I will then in the discussion section try to propose a set of research questions for CSCW researchers who investigate digital labor platforms.

Theoretical Background

In this section, I give a short overview of how I believe the traditional view of collaboration developed in CSCW has been commoditized in digital labor platforms. I then discuss how platform owners facilitate this commoditization, mainly through their governance models. I conclude with a short description of platform boundary resources as one way to structure future research and impact.

CSCW research during the last decades has played a central role in the emergence of today's digital labor platforms. CSCW, through its rigorous studies of work practices, succeeded in creating an understanding, and partly codifying collaboration into various theories (Schmidt & Bannon, 2013). CSCW researchers created knowledge about collaboration, and how it can be supported across time and space using IT-based tools. Distributed coordination mechanisms were demonstrated in tools such as Ariadne (Simone & Divitini, 1998) and later made commercially available –albeit in modified versions –in workflow tools. Elements from the speech act theory (Medina-Mora, Winograd, Flores, & Flores, 1993) were incorporated in commercial messaging systems. Various systems were developed

to support situated action based on, e.g. awareness (Dourish & Bellotti, 1992; Gross, 2013) and so on. These experiments resulted in products and features that we take for granted today: Awareness information about our friends and colleagues is now everywhere (sometimes also too much of it!); various systems implement more or less flexible workflows guiding (or forcing) us to get the work done.

This understanding and codification of collaboration into theories and digital tools was a prerequisite for a full digitalization of collaboration in the form of platforms. An early prototype, the BSCW system (Bentley, Horstmann, & Trevor, 1997) is an illustrative example of how two basic concepts originating from CSCW, i.e., common information spaces (Bannon & Bødker, 1997) and awareness mechanisms (Dourish & Bellotti, 1992) were used to support any document-based collaboration across time and space and without any physical contact among collaborators. The ability to cross the boundaries of time, space and organizations –partly fueled by research from CSCW –has enabled platforms to go from being internal and isolated tools for individual organizations to become open platforms for global industries (Gawer, 2014). They have created a vocabulary known to a global workforce. Everybody knows now what an "Uber", a "Google doc" or "retweets" or "likes" or "feed updates" etc. are[1].

Standardization often goes together with commoditization. Commodity was discussed and defined by Marx: "Hence, commodities are first of all simply to be considered as *values*, independent of their exchange-relationship or from the *form*, in which they *appear as exchange-values*" (Marx, 1867). Wikipedia defines commodity as "an economic good or service that has full or substantial fungibility: that is, the market treats instances of the good as equivalent or nearly so with no regard to who produced them." In this extremely short review of definitions, three properties of a commodity appear to be central. First, commodities are created to be exchanged, so they need to be packaged. Second, commodities have values, so that they can be traded. Third, commodities are fungible, so a commodity from one source can be replaced by one from another. But can such an intellectually loaded activity such as collaboration become commoditized? My answer is yes. Platforms commoditize collaboration along at least two lines. First, they standardize collaboration by simplifying it and eliminating its contextual dependencies –i.e., packaging collaboration as goods. Second, they use various mechanisms to create and grow a market for trading "packaged collaboration." I will shortly discuss each of these aspects.

Standardization and packaging of collaboration

A growing number of publications in CSCW and HCI already show us how platforms are simplifying and decontextualizing collaboration. The extreme

[1] This standardization is helped by the fact that CSCW often tends to be agnostic about who is collaborating with whom, and focuses instead on their actions and "embodiment" in a virtual world.

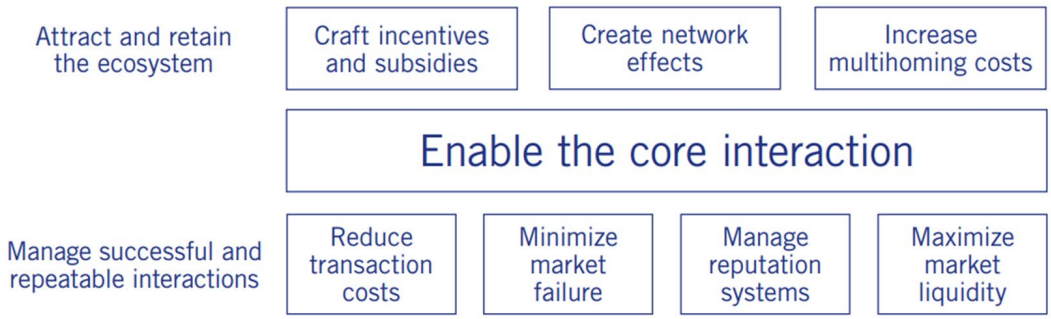


Figure 2: Mechanisms used by digital labor platforms to create and grow a marketplace for collaboration (Choudary, 2018).

examples come from micro-task platforms such as Amazon Mechanical Turk. Micro-task platforms, as the name suggests, break down the tasks into small pieces: "Turkers (termed 'Providers' by AMT) are the users completing the [Human Intelligence Tasks], which typically take seconds or minutes paid at a few cents at a time" (Martin, Hanrahan, O'Neill, & Gupta, 2014). This breaking down of tasks helps eliminate the need for specialized skills: "Such simple, small-scale work has engendered low-pay, piece rate reward structures, in part due to the perception that workers are homogenous and unskilled" (Kittur et al., 2013). Kittur et al. also argue that crowdsourcing platforms fail to support more –intellectually –complex tasks and workflows: "The current model is... insufficient to support the complexity, creativity, and skills that are needed for many kinds of professional work that take place today. Nor can it drive factors that will lead to increased worker satisfaction, such as improved pay, skill development, and complex work structures" (ibid, p. 1303).

In addition to breaking down and simplifying tasks –thereby increasing fungibility and facilitating trade –platforms replace the need for local skills and knowledge with less specialized or completely new standardized skills –often in favor of consumers. A much-discussed example is the new skills of "emotional labor" (Raval & Dourish, 2016) that are increasingly required from platform workers. Uber drivers, for instance, are not anymore required to have "the knowledge" –of all the local streets and addresses –but use GPS guides and instead engage in standardized emotional labor: "skills of engaging with passengers shape the self- image of the Uber driver" (Glöss et al., 2016, p. 1637). In this way, standardized packages of exchangeable labor are created that allow platforms to be deployed in different settings and cultures without the need for any specialized or local knowledge.

Marketplace mechanisms

Once the collaboration is "packaged", it needs to be sold. So, there is a need for a marketplace. Labor platforms do not only standardize and support collaboration

between consumers and workers. They also use several mechanisms to create and grow a marketplace for collaboration. Some of the mechanisms that operate in such a marketplace are excellently described by Choudary (Choudary, 2018) and shown in Figure 2. The core interaction –the large central box in the figure—depicts the support for collaboration between consumer and worker –what has traditionally been of interest for CSCW. The boxes surrounding the core interaction are the mechanisms often implemented by platform owners to grow the marketplace.

The top row of mechanisms is aimed at attracting and retaining collaborators. Digital labor platforms –as many two-sided economies –create network effects to attract new users. For instance, the more workers with a good reputation you have on your platform, the more consumers and workers will want to use the platform. Platforms also use incentive systems to attract new users. Many labor platforms subsidize consumers at the expense of workers because consumers who are willing to pay will bring new workers to the platform. Platforms use various mechanisms to make it difficult for users to move to other platforms or operate on multiple platforms (multihoming). For instance, building a reputation through e.g. star ratings (Wilson & Paoli, 2018) is a demanding task for workers. Platforms often take ownership of a worker's star ratings and make it impossible to move reputation to other platforms if a worker wanted to do so.

The bottom row in Figure 2 depicts activities that are used to alter the core interaction in order to grow the platform. These are activities that have greatest impact on the core interaction and thereby of high relevance for CSCW. Reducing transaction costs is done partly by simplifying tasks as discussed above. Additionally, platforms use automated matchmaking, with the consequence that collaborators on a labor platform rarely know or see each other: "the task creator [in Amazon Mechanical Turk] has no way of knowing if the task worker is male or female, young or old, religious or atheist, etc." (Gray et al., 2016, p. 134). Researchers have shown that there is an asymmetry in access to information in platforms. Normally, users know very little about each other, and platform owners know much more about users: "[Amazon Mechanical Turk] is something of a 'black box.' That is, while Amazon does publish their terms and conditions, little information is released about how these policies are specifically realised " (Gupta et al., 2014). This lack of transparency can often result in decreased quality in the collaboration between workers and consumers (Kittur et al., 2013). The use of data-driven algorithms is pointed out as a contributor to information asymmetry and imbalance of power between Uber and its drivers (Rosenblat & Stark, 2016). Reputation systems, such as star ratings, intend to control the quality of the provided services but can do the opposite because of the sanctions posed on platform users that strongly affect the interactions among them: "The drivers are scared of the customers but also the customers are scared of the drivers" (Glöss et al., 2016, p. 1635). Through automated matchmaking and global competition,

platforms also try to exploit fungibility and bring the price of labor to a minimum (Martin et al., 2014).

Co-creating platform governance models

Standardization, simplification, and packaging of interactions, and the additional activities of creating a marketplace for labor are not done in a vacuum. They are often parts of orchestrated activities to increase revenue and market share for platform owners. These activities are often aimed at creating so-called *platform governance models*. The way a platform uses labor standardization, subsidies, network mechanisms etc., as discussed above, constitutes that platform's governance model. Governance models "would let platforms control interactions between multiple stakeholders without jeopardizing their incentives for value-creation" (Constantinides et al., 2018, p. 383). Platform governance models have emerged as a major research topic in the information systems research field. Most research on platform governance models takes the perspective of the platform owner (Schreieck, Wiesche, & Krcmar, 2016). However, as I hope I have demonstrated above, the impact of governance models on collaboration, and therefore on CSCW, can be very real.

Platform governance models can be an instrument for CSCW researchers to increase our impact. Governance models have the advantage of taking an ecosystem perspective and avoid focusing on only one or a few actors. Emerging research in information systems shows how governance models can be co-created, as pointed out by Schreieck et al.: "Including the complementors and end-users into the analysis, will also allow to discuss a bottom-up approach in the design and governance instead of interpreting it as a top-down approach only" (2016). By leveraging CSCW research on work practices, we increase the chance of influencing the design and evolution of governance models. We have argued elsewhere (Paper submitted to ECSCW 2019) how governance models can be co-created by using IT-based *boundary resources* that platform owners implement in order to enable interactions with their platforms (Ghazawneh & Henfridsson, 2013). The model of platform boundary resources can be useful for CSCW research because of several reasons. First, as shown in the literature, e.g. (Eaton, Elaluf-Calderwood, Sorensen, & Yoo, 2015; Islind, Lindroth, Snis, & Sørensen, 2016), it gives us an analytical tool to connect digital workplace studies to the study of platform governance models. In this way, it creates a bridge for dialog and a point of impact for CSCW. Second, CSCW researchers are already familiar with the concept of boundary resources (Leigh Star, 2010). This knowledge can be used to make efficient use of the model and create better and fairer platform designs.

Discussion

In this paper, I have aimed to demonstrate how digital labor platforms build on the notion of collaboration as commodity, and how platform governance models are used to simplify, standardize, package and trade collaboration as commodities in a marketplace. I have argued that CSCW needs to reconsider its impact on how these platforms are developed. CSCW research on digital labor platforms, in order to have lasting impact, needs to construct new, viable –bottom-up-driven – alternatives to existing platform governance models. This implies addressing several research questions that I aim to emphasize in this discussion section.

In dealing with the emerging landscape of digital labor platforms, CSCW researchers have done what they are good at, i.e., studying how people use these digital tools to cooperate. Digital workplace studies have provided us with crucial knowledge about users' interactions with, and opinions of, platforms. Although such digital workplace studies are important, CSCW researchers also need to focus on the whole ecosystem of stakeholders and not only workers. Finding a balance between the values held by workers, consumers and platform owners can result in increased impact. This impact is sorely needed in order to redirect the development of platforms and include more of the needs and preferences of the users.

A fundamental first research stream is needed in order to understand the consequences of packaging and decontextualizing collaboration. At the core of CSCW is the fact that collaboration is local and contextual (Suchman, 1987). The research question is "What are the consequences of packaging and decontextualization of labor for workers and consumers?" We already see emerging research looking into the challenges that simplification of labor introduces both for workers, in terms of workplace quality, and for consumers, in terms of the quality of the results they get (Kittur et al., 2013). It is important to demonstrate through research how workers and consumers are affected –positively or negatively by platforms.

Second, we need to look at how commoditization activities –subsidizing, increasing multihoming costs, etc., as discussed above –impact collaboration and its results. Our well-established theories about IT-based collaboration might be affected by the fact that a new –and strong –agent, i.e. the platform owner, has entered the stage and is willing to impose changes on how we collaborate. An important research question might then be: "How do platform owners and their governance models affect collaboration among consumers and workers? And what is the impact on the quality of the results?" Several referenced studies in this paper try to answer this question. We need to further look into what commoditization of collaboration means, i.e., when buying and selling collaboration becomes the main focus –instead of how collaboration is done –how will that affect our research? A research stream might be to investigate: "What is the perceived business value of collaboration for its beneficiaries?"

Third, we need to start addressing a new set of research questions regarding the collaboration between platform owners and platform users. We already know something from existing research about the nature of this collaboration. But in general, most CSCW research is about users only, as most IS research is about platform owners only. We need to continue generating more knowledge in order to be able to address the research question: "How does the collaboration between platform owners and platform users happen? What are the underlying values for each collaborating party? And how can the goals of this type of collaboration be achieved?" Moreover, we need to know how this collaboration can be supported digitally. Therefore, another research question for CSCW should be: "How do we make collaboration happen between platform owners and users? What arenas – digital or offline –do we need for this type of collaboration?" One particular area in need of research is how we can replace off-line arenas for discussing working conditions and wages with online arenas, as noted by Glöss et al.: " [Amazon Mechanical Turk] and Airbnb return labour issues to relevance, since the apps are involved in payment income, rates, productivity and conditions of the work being completed through them" (Glöss et al., 2016).

A more fundamental question in my view is related to platform governance models. As CSCW researchers, we need to know more about these models and find ways of influencing them with our knowledge. Our knowledge needs to be combined with other types of knowledge from management, economics, market regulation, labor unions, etc. A better workplace for workers cannot be created if it is not economically viable for the platform owner, if it is not manageable, or if it is not regulated by laws. So, the research question to address is: "How can CSCW researchers, together with researchers from other disciplines, help co-create digital labor platform governance models that are fair?"

Conclusions

In this paper, I have discussed the role of CSCW in the landscape of digital labor platforms. My argument has been that CSCW needs to pay more attention to the underlying governance models of these platforms. I have argued that we need to have new models and tools that allow us to co-create these governance models. I use the model of Platform boundary resources as an example of how such co-creation can be done practically[2].

Our future research in this direction includes a thorough analysis of existing CSCW and HCI research with the lenses of governance models. One goal is to construct design guidelines for boundary resources based on our knowledge

[2] See also our case study of small-scale platform co-creation, to be presented at ECSCW 2019: Farshchian, B.A., Thomassen, H.E. (2019 forthcoming): *Co-creating platform governance models using boundary resources: A case study from dementia care services*.

contained in workplace studies. Such guidelines can be used as a tool for dialog among the different stakeholders in the digital labor ecosystems.

Confronted with the global reach and market size of some of the largest global labor platforms, it is easy to doubt that our research can have an impact on platform governance models. My view is that CSCW has a lot to offer and can act as a force for creating alternative realities in the field of platform governance models. Such alternative realities have a big chance of being both fair and sustainable at the same time, and in this way create our future global labor platform models.

Acknowledgement

This research was partly funded by the Norwegian Research Counsel's FORKOMMUNE program, project SMED. I thank anonymous reviewers of ECSCW 2019 exploratory papers for useful feedback on an earlier draft. I also thank the participants of Forskerfabrikken for feedback and Eric Monteiro for coordinating the "Fabrikken." I thank Marius Mikalsen, Jacqueline Floch, Thomas Vilarinho, and Hanne Ekran Thomassen for fruitful discussions regarding platforms, some of which led to the idea of this paper.

References

Ahmed, F., & Fuge, M. (2017). Capturing Winning Ideas in Online Design Communities. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17* (pp. 1675–1687). New York, New York, USA: ACM Press. https://doi.org/10.1145/2998181.2998249

Bannon, L., & Bødker, S. (1997). Constructing Common Information Spaces. In *Proceedings of the Fifth European Conference on Computer Supported Cooperative Work* (pp. 81–96). Berlin, Heidelberg: Springer-Verlag. https://doi.org/10.1007/978-94-015-7372-6_6

Bentley, R., Horstmann, T., & Trevor, J. (1997). The World Wide Web as Enabling Technology for CSCW: The Case of BSCW. *Computer Supported Cooperative Work (CSCW)*, *6*(2–3), 111–134. https://doi.org/10.1023/A:1008631823217

Bratteteig, T., & Wagner, I. (2016). Unpacking the Notion of Participation in Participatory Design. *Computer Supported Cooperative Work (CSCW)*, *25*(6), 425–475. https://doi.org/10.1007/s10606-016-9259-4

Choudary, S. P. (2018). *The architecture of digital labour platforms: Policy recommendations on platform design for worker well-being*. International Labour Organization.

Constantinides, P., Henfridsson, O., & Parker, G. G. (2018). Introduction—Platforms and Infrastructures in the Digital Age. *Information Systems Research*, *29*(2), 381–400. https://doi.org/10.1287/isre.2018.0794

de Boer, P. M., & Bernstein, A. (2017). Efficiently Identifying a Well-Performing Crowd Process for a Given Problem. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social*

Computing - CSCW '17 (pp. 1688–1699). New York, New York, USA: ACM Press. https://doi.org/10.1145/2998181.2998263
de Reuver, M., Sørensen, C., & Basole, R. C. (2018). The Digital Platform: A Research Agenda. *Journal of Information Technology*, *33*(2), 124–135. https://doi.org/10.1057/s41265-016-0033-3
Dourish, P., & Bellotti, V. (1992). Awareness and coordination in shared workspaces. In *Proceedings of the 1992 ACM conference on Computer-supported cooperative work - CSCW '92* (pp. 107–114). New York, New York, USA: ACM Press. https://doi.org/10.1145/143457.143468
Eaton, B., Elaluf-Calderwood, S., Sorensen, C., & Yoo, O. (2015). Distributed Tuning of Boundary Resources: The Case of Apple's iOS Service System. *MIS Quarterly*, *39*(1), 217–243.
Frenken, K., & Schor, J. (2017). Putting the sharing economy into perspective. *Environmental Innovation and Societal Transitions*, *23*, 1–8. https://doi.org/10.1016/j.eist.2017.01.003
Gawer, A. (2014). Bridging differing perspectives on technological platforms: Toward an integrative framework. *Research Policy*, *43*(7), 1239–1249. https://doi.org/10.1016/j.respol.2014.03.006
Ghazawneh, A., & Henfridsson, O. (2013). Balancing platform control and external contribution in third-party development: The boundary resources model. *Information Systems Journal*, *23*(2), 173–192. https://doi.org/10.1111/j.1365-2575.2012.00406.x
Glöss, M., McGregor, M., & Brown, B. (2016). Designing for Labour: Uber and the On-Demand Mobile Workforce. In *CHI '16. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, California, USA, May 7-12, 2016* (pp. 1632–1643). New York, New York, USA: ACM Press. https://doi.org/10.1145/2858036.2858476
Gray, M. L., Suri, S., Ali, S. S., & Kulkarni, D. (2016). The Crowd is a Collaborative Network. In *CSCW '16. Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing, San Francisco, California, USA, February 27-March 2, 2016* (pp. 134–147). New York: ACM Press. https://doi.org/10.1145/2818048.2819942
Gross, T. (2013). Supporting effortless coordination: 25 years of awareness research. *Computer Supported Cooperative Work: CSCW: An International Journal*, *22*(4–6), 425–474. https://doi.org/10.1007/s10606-013-9190-x
Gupta, N., Martin, D., Hanrahan, B. V., & O'Neill, J. (2014). Turk-Life in India. In *GROUP'14. Proceedings of the 18th International Conference on Supporting Group Work, Sanibel Island, Florida, USA, November 9-12, 2014* (pp. 1–11). New York: ACM Press. https://doi.org/10.1145/2660398.2660403
Harmon, E., & Silberman, M. S. (2018). Rating Working Conditions on Digital Labor Platforms. *Computer Supported Cooperative Work (CSCW)*, *27*(3–6), 1275–1324. https://doi.org/10.1007/s10606-018-9313-5
Hata, K., Krishna, R., Li, F.-F., & Bernstein, M. S. (2017). A Glimpse Far into the Future: Understanding Long-term Crowd Worker Quality. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17* (pp. 889–901). New York, New York, USA: ACM Press. https://doi.org/10.1145/2998181.2998248
Huws, U., Spencer, N. H., & Joyce, S. (2016). *Crowd Work in Europe - Preliminary results from a survey in the UK, Sweden, Germany, Austria and the Netherlands*. *FEPS report* (Vol. 58). Retrieved from https://www.feps-europe.eu/resources/publications/453-crowd-work-in-europe.html

Islind, A. S., Lindroth, T., Snis, U. L., & Sørensen, C. (2016). Co-creation and Fine-Tuning of Boundary Resources in Small-Scale Platformization. In L. S. U. (Ed.), *Nordic Contributions in IS Research. SCIS 2016. Lecture Notes in Business Information Processing* (Vol. 259, pp. 149–162). Springer. https://doi.org/10.1007/978-3-319-43597-8_11

Kittur, A., Nickerson, J. V., Bernstein, M., Gerber, E., Shaw, A., Zimmerman, J., … Horton, J. (2013). The future of crowd work. In *CSCW'13. Proceedings of the 2013 conference on Computer supported cooperative work, San Antonio, Texas, USA, February 23-27, 2013* (pp. 1301–1318). New York: ACM Press. https://doi.org/10.1145/2441776.2441923

Leigh Star, S. (2010). This is Not a Boundary Object: Reflections on the Origin of a Concept. *Science, Technology & Human Values*, *35*(5), 601–617. https://doi.org/10.1177/0162243910377624

Martin, D., Hanrahan, B. V, O'Neill, J., & Gupta, N. (2014). Being a turker. In *CSCW'14. Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing, Baltimore, Maryland, USA, February 15-19, 2014* (pp. 224–235). New York: ACM Press. https://doi.org/10.1145/2531602.2531663

Marx, K. (1867). *Capital* (1st ed.).

Medina-Mora, R., Winograd, T., Flores, R., & Flores, F. (1993). The action workflow approach to workflow management technology. *The Information Society*, *9*(4), 391–404. https://doi.org/10.1080/01972243.1993.9960152

Raval, N., & Dourish, P. (2016). Standing Out from the Crowd: Emotional Labor, Body Labor, and Temporal Labor in Ridesharing. In *CSCW'16. Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing, San Francisco, California, USA, February 27- March 2, 2016* (pp. 97–107). New York: ACM Press. https://doi.org/10.1145/2818048.2820026

Rodes, A. (2017). Uber: Which countries have banned the controversial taxi app. *Independent*. Retrieved from https://www.independent.co.uk/travel/news-and-advice/uber-ban-countries-where-world-taxi-app-europe-taxi-us-states-china-asia-legal-a7707436.html

Rosenblat, A., & Stark, L. (2016). Uber's Drivers: Information Asymmetries and Control in Dynamic Work. *International Journal of Communication*, *10*(27), 3758–3784. https://doi.org/10.2139/ssrn.2686227

Schmidt, K., & Bannon, L. (2013). Constructing CSCW: The first quarter century. *Computer Supported Cooperative Work: CSCW: An International Journal*, *22*(4–6), 345–372. https://doi.org/10.1007/s10606-013-9193-7

Schreieck, M., Wiesche, M., & Krcmar, H. (2016). Design and Governance of Platform Ecosystems – Key Concepts and Issues for Future Research. In *ECIS 2016. Twenty-Fourth European Conference on Information Systems, İstanbul, Turkey, June 12-15, 2016* (pp. 1–20). AIS.

Semuels, A. (2018). The Internet Is Enabling a New Kind of Poorly Paid Hell. *The Atlantic*. Retrieved from https://www.theatlantic.com/business/archive/2018/01/amazon-mechanical-turk/551192/

Simone, C., & Divitini, M. (1998). Ariadne: Supporting Coordination Through a Flexible Use of Knowledge Processes. In *Information Technology for Knowledge Management* (pp. 121–148). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-03723-2_7

Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine communication.* Cambridge University Press.

Tiwana, A. (2013). *Platform Ecosystems: Aligning Architecture, Governance,*

and Strategy. Waltham, MA, USA: Morgan Kaufmann Publishers Inc.
Wilson, A., & Paoli, S. De. (2018). Reconsidering online reputation systems. In *Proceedings of 16th European Conference on Computer-Supported Cooperative Work - Exploratory Papers* (pp. 1–18). Nancy, France: European Society for Socially Embedded Technologies (EUSSET). https://doi.org/10.18420/ecscw2018_12

Xinru Tang; Dongyang Zhao; Ying Zhang; Xianghua Ding (2019): AuDi: an Auto-Feedback Display for Crowdsourcing. In: Proceedings of the 17th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing an the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep05

AuDi: an Auto-Feedback Display for Crowdsourcing

Xinru Tang, Dongyang Zhao, Ying Zhang, Xianghua Ding
Shanghai Key Laboratory of Data Science, Fudan University; School of Computer Science, Fudan University
{15300160005,15307110394,16210240023,dingx}@fudan.edu.cn

Abstract. While feedback, by experts or peers, is found to have positive effects on crowdsourcing work, it is a costly approach as more people or time is involved in order to provide feedback. This paper explores an automatic feedback display called AuDi for crowdsourcing. AuDi shows the worker's accuracy rate, which is automatically calculated with the use of an accuracy algorithm, by changing the background color of the task page. We conducted an experimental study with AuDi in the field, and employed both quantitative and qualitative methods for data collection and analysis. Our study shows that, without introducing new cost, such an auto-feedback display is well received by our participants, gives them assurance and more confidence, and also positively contributes to work performance by pushing them to study more and understand better the task requirements.

1 Introduction

Work performance – particularly in terms of quality output, and work experience are common concerns for crowdsourcing. Many factors could lead to quality issues in crowdsourcing, including unqualified workers (Rzeszotarski and Kittur, 2011; Gadiraju et al., 2015), misunderstanding of requirements (McInnis et al., 2016; Kulkarni et al., 2012; Ipeirotis et al., 2010), and so on. A variety of quality control mechanisms have been explored, such as redundancy and majority voting (Callison-Burch, 2009; Franklin et al., 2011), adding test questions to obtain

accuracy of the workers' answer (Liu et al., 2012), using algorithms to infer the true answer such as Bayesian theory or Expectation Maximization (Liu et al., 2012; Ipeirotis et al., 2010) and etc. Some platforms such as AMT simply reject unqualified work after all the tasks are completed, which, however, causes a series of negative effects on workers' experience (McInnis et al., 2016).

In recent years, feedback as a way to enhance crowd work experience and improve quality output has been investigated (McInnis et al., 2016; Dow et al., 2012). Research has shown positive effects of feedback for crowd work. For instance, a study employing self-assessment and expert reviews as feedback illustrates that, "timely, task-specific feedback helps crowd workers earn, persevere, and produce better results" (Dow et al., 2012). While generally positive, however, most of these are based on personal feedback, which may cause an increase of cost as it relies on more people to spend time on giving feedback.

In this paper, we are investigating an approach which provides automatic feedback to the crowd workers in a timely fashion. More specifically, an accuracy algorithm, based on an accuracy calculation method (Feng et al., 2014), is employed and its accuracy result is shown as the background color of the task page in real time as an ambient feedback display. We refer to this ambient automatic feedback display as AuDi in this paper.

With the study, we found that AuDi was positively perceived and well taken into their crowd work. Both qualitative and quantitative results show that AuDi enables participants to know better of their own performance, feel more in control, and enhance their confidence.

2 Related Work

2.1 Quality Control in Crowdsourcing

Crowdsourcing relies on workers' good performance to produce high-quality output. However, since workers involved are from different countries, with different ages and educational levels, their subjective awareness and background knowledge would inevitably affect their understanding and interpretation of task requirements(Ross et al., 2010; Martin et al., 2014; Gadiraju et al., 2015; Ipeirotis et al., 2010). In consequence, the output quality is barely satisfactory, leading quality evaluation and control to be big issues in crowdsourcing (Kittur et al., 2013).

Many algorithms are proposed to measure the quality of submitted answers. The most common method is redundancy and majority voting, in which the answer given by the majority workers is taken as the correct answer (Callison-Burch, 2009; Franklin et al., 2011; Kulkarni et al., 2012; Little et al., 2010). Further, redundancy can not only be used to determine the correct answer but also help to evaluate the accuracy of each worker (Ipeirotis et al., 2010).

Some research has applied workers' accuracy to the estimation of the results, by integrating workers' answers and their accuracy to infer the correct answer. For

example, one strategy is based on the Expectation Maximization (EM) algorithm, which calculates workers' accuracy by using the confusion matrix (Ipeirotis et al., 2010). This method obtains high-quality results but is at the expense of long inference time.

Besides these underlying quality control algorithms, mechanisms are also explored to change the workflow as a way to enhance work performance for quality output. For instance, Wiseman et al. experimented with inserting an additional check stage, however, their results showed that this would not reduce the error rate, because only a check stage does not make people bother to check their answers (Wiseman et al., 2013). Sandy J. J. Gould et al. studied the effect of a lockout in a data-entry task, and similarly the research shows that the lockout mechanism does work in a laboratory setting, but not in the field where people will do other tasks during the lockout period, making lockouts no longer effective (Gould et al., 2016).

2.2 Work Experience in Crowdsourcing

In recent years, not simply quality output, but the quality of work experience has also become a concern for crowd work. As mentioned, rejecting unqualified work is commonly adopted for quality control, however, work is usually rejected by the requester without giving reasons. This is problematic since payment is the primary motivation of workers (Janine, 2016; Brewer et al., 2016). Past research showed that many workers reported not being paid for adequately completed tasks (McInnis et al., 2016; Irani and Silberman, 2013). Users express their concern about submitting unqualified work, and they are also worried that they may not understand the task which would lead to the failure to get the pay (Mao et al., 2013). This also leads to general feelings of unfairness around rejection, since the requester can get access to all the information about the user, while the users know nothing about their performance and the job criteria (McInnis et al., 2016).

To improve crowd work performance and experience, some particularly focus on providing feedback in real time. For example, Dow et al. (Dow et al., 2012) studied different feedback mechanisms, including peer review, expert review and self-assessment, and found that both self-assessment and feedback from outside will significantly increase the work quality. Concerned with risks in user experience caused by reasons such as unclear evaluation criteria, Brian McInnis et al. (McInnis et al., 2016) suggested automated feedback, which will enable the user to know their performance in time, so it can help build trust between the users and requesters, protect honest users from honest error, and meanwhile punish bad actors.

3 Method

Our study was based on a crowdsourcing platform named ZhongYan, which was set up by our lab for crowd work research projects. For the study, we carefully chose

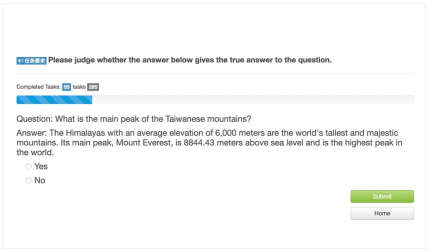


Figure 1. An example of the tasks.

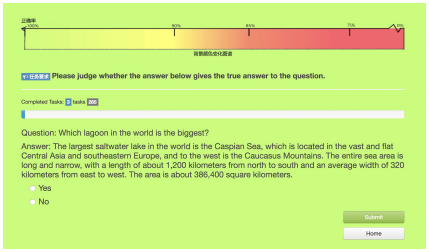


Figure 2. An example for the experimental interface.

and designed crowd task, accuracy algorithm, feedback display format, as well as the experiments, which we will elaborate below.

3.1 Task

A set of text annotation tasks from a real-world project was chosen for our study. For each task, workers are shown one question along with one answer and their job is to determine whether the correct answer appears in the answer text by choosing "yes" or "no" on the task page. We chose it because it was representative of typical crowd task, and was of medium difficulty, which means some may come across a question beyond his or her knowledge, yet he or she can get over this by making extra efforts, for example, Googling. Figure 1 shows an example of the task. In the experiment, every participant was asked to finish about 250 tasks.

3.2 Accuracy Algorithm

For accuracy algorithm, we chose and adapted a quality evaluation algorithm of crowd work proposed by J. Feng et al. (Feng et al., 2014). It is based on Majority Vote (MV), a very popular method to infer the final results in crowdsourcing, and further improves the inference results by considering the different qualities for each worker. This method was chosen because it can achieve a good balance between calculation accuracy and response time compared to other methods (Raykar et al., 2010; Ipeirotis et al., 2010), as it uses an incremental rather than iterative strategy to update the workers' quality. Two models are used in this incremental algorithm. One is the worker model and the other is the question model. The worker model is a quadruple and each element in the quadruple is presented as c_{ij}(i means the answer given by the worker and j means the true answer to the question).

$$\begin{bmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{bmatrix}$$

And the accuracy rate of each worker is calculated as:

$$acc = \frac{c_{00} + c_{11}}{c_{00} + c_{01} + c_{10} + c_{11}}$$

While the question model is presented as a tuple (p_i, 1-p_i) in which p_i is the probability that the true answer to the question is the first choice. And if p_i > 1- p_i, it takes the first choice as the true answer. We build the worker model for each worker to compute the accuracy of the worker and build the question model for each question to infer its result. Each time a worker submits his/her answer we will incrementally update the question model, and the worker model will be updated when we decide the answer to the question in order to acquire the worker's accuracy timely. For our experiments in particular, we used twenty test questions with which we know the correct answers to initialize the worker model of each worker. When it comes to the official questions which lack the correct answers, we compute the question model for each question in order to infer the correct answer combining the submitted answer and the submitter' accuracy computed by his worker model. And we updated each worker's worker model according to the calculated result in order to compute the worker's accuracy timely.

3.3 The Auto-Feedback Display

We decided to show the accuracy information in an ambient form (Mankoff et al., 2003), as it is suitable for persuasion without obtrusion. More so, as representing feedback via color of surrounding area is found to be easier to process and use in goal-striving processes than factual feedback (Ham and Midden, 2010), we decided to use the background color of the task page as a way to show the accuracy information.

As such, we implemented a display schema altering background color of the task page based on accuracy. The color schema is inspired from traffic lights, with red standing for dangerous status, yellow for warning and green for safety. Every time a worker submits his or her answer, the website will change its background color according to the newly calculated accuracy rate while loading the task page.

Through a pilot study on the tasks, we correspond the color schema with an accuracy range from 75% to 100%, so that participants could easily see the change of color while working on the tasks. That is, the accuracy rate of 75% corresponds to the reddest color, and the accuracy rate of 100% corresponds to the greenest. For the sake of convenience, this color schema was shown as a bar on the task page, shown in Figure 2 to help people understand the meaning of the background color.

We chose an accuracy rate of 80% as the acceptance rate for the task. To simulate the real world situation, the participants were told that only those who completed all tasks with an accuracy rate over 80% would earn 50 RMB, otherwise, they would not get paid. Besides, they were also told that they could terminate the experiment anytime they want but only those who finish it in the scheduled time would get paid. After the end of the experiment, however, all got paid as a compensation for their participation in the study.

3.4 Experiment

A total of 50 participants were recruited for the study. These participants were evenly divided into control group and experiment group. Some participants did not show up during our experiments, so at last, there were 22 participants in the control group, 22 in experiment group. We gave participants two days to complete all the tasks, so they could choose any time they like and pause whenever they want, as a way to simulate the real world situation.

Workers in the control groups did not have any feedback - that is, the web background color stayed white (Figure 1), while workers in the experiment group was provided with our automatically calculated accuracy rate as feedback as mentioned above (Figure 2). Participants in the experiment groups were informed of the basic idea of accuracy algorithm in use, and the feedback display. To make it closer to a real-world project, we adopted a redundancy of 5. That is, we divided workers in the experiment groups to subgroups of 5 to calculate accuracy rate within each subgroup. The back-end of the platform recorded each participant's answers and work time for later analysis.

After they finished the project, all workers were assigned an online questionnaire the minute they finished the project to report their self-assessment and personal experience. Participants in the experiment groups were also asked to answer questions about their experience regarding the feedback display while the control groups didn't need to. Almost all the questions of the questionnaire were given in Likert 5-point, except for one question asking participants to write their expected accuracy. Questions covered concentration, confidence, expected accuracy, perseverance and so on, and these data would be for quantitative analysis.

We also conducted interviews with the experiment groups. We recruited 10 interviewees before the experiment started and sought out one more who quit after doing 14 tasks after the experiment. Detail information of these interviewees are listed in Table I ('P' denotes the participants).

ID	Accuracy Rate(%)	ID	Accuracy Rate(%)
P1	86.7	P2	93.3
P3	92.6	P4	89.8
P5	88.8	P6	87.0
P7	88.8	P8	91.9
P9	93.0	P10	90.2
P11	78.6		

Table I. Interviewees' Information.

All the interviews were conducted online through text chat. Each interview lasted about 30 minutes, in which each interviewee reported their experiences about the feedback display and any trouble they came across during the experiment. During our interviews, our questions mainly focused on how they felt

about the feedback display, whether it accorded with their own estimation, how they were possibly influenced by the feedback display, and whether they liked to have the display or not, and why.

3.5 Data Analysis

For quantitative analysis, we collected data from the back-end of the platform and calculated drop out rate, pass rate and accuracy rate. Specifically, we calculated mean, median and standard deviation of these measures for comparison. Meanwhile, we collected results from the questionnaire regarding their estimated accuracy, their confidence level, and so on. For qualitative analysis, we went through the interview data, and identified themes emerged from it. We paid particular attention to those themes that are related to the quantitative results we found from the analysis.

4 Results

Overall, almost all our participants from the experiment groups perceived the feedback provided by AuDi as well reflected their performance, and would all preferred to have it for their crowd work. They further reported that the feedback display had positively influenced their performance and experience, e.g. the red color made them pause for thought and the green color encouraged them to continue. In this way, participants adjusted their work pace accordingly. Below, we will present our results of how people perceived the feedback provided by AuDi, and how AuDi had effects on their performance and experience of doing crowd work.

Group	Completed	Dropped out	Total
Experiment Group	19	3	22
Control Group	21	1	22
Total	40	4	44

Table II. Drop Out Rate.

Group	Passed	Failed	Total
Experiment Group	19	0	19
Control Group	20	1	21
Total	39	1	40

Table III. Pass Rate.

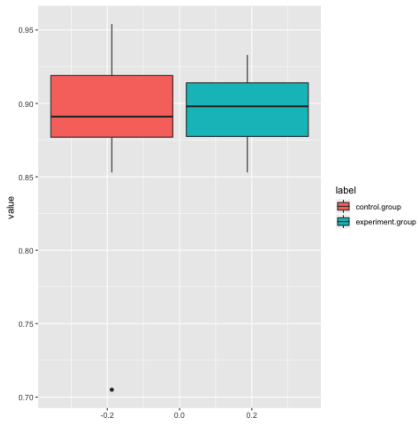


Figure 3. Accuracy Rate.

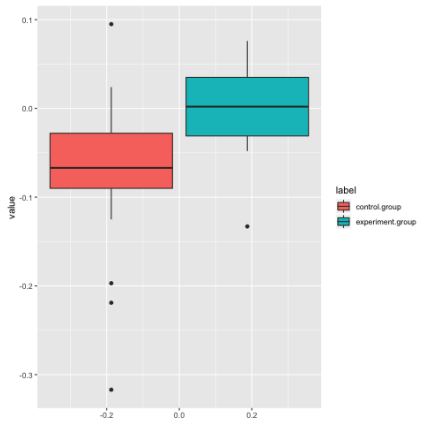


Figure 4. Difference between Self-Estimated and Real Performance.

4.1 Feedback Accuracy and Acceptability

As for the accuracy of the feedback itself, our data shows that participants took the feedback provided by AuDi as largely reliable and quite acceptable. According to the results from the questionnaire, the majority of participants considered the feedback was consistent with their own estimations. Specifically, 26.83% chose level 3, 51.22% level 4, and 9.76% level 5. Our interview data further shows that they considered the feedback was in an acceptable range. For example, P8 commented:

> "Since 100% reliability is unrealistic, I just need feedback stable enough to help me develop a general sense of direction. I mean, the more information provided, the more helpful."

4.2 Effects on Work Performance

Overall, we found the experiment groups performed better and more stably (with less fluctuation of accuracy rate) than the control groups (control group SD_1=0.048, experiment group SD_2=0.029). As shown in Table III, the only one who completed the experiment but did not meet the accuracy rate bar (80%) was from the control group. In addition, Figure 3 shows that the median accuracy rate of the experiment group is higher than the control groups (M_1=0.891, M_2=0.898). Besides, it also shows that the only outlier was from the control group: it was 0.748.

Our interview analysis further suggests that the performance difference between experiment and control groups had something to do with the different levels of understanding of the task requirement, with or without AuDi. Many of our participants reported that they did not understand what they were asked to do, although the description and requirement of the task had been given before the experiment, and few people would bother to read the long description of requirements carefully before the task. The use of AuDi pushed participants to

reread the requirement descriptions when they got relatively negative feedback, which, to an extent, enhanced their understanding and so was their performance.

In our interviews, participants from the experiment group reported how AuDi helped them to learn and grasp the knack to solve the task in practice. P3 described to us:

> "At first, I did not figure out the true requirement of this task. However, the screen changed red abruptly after I gave the wrong answers. Then, I read the question and text again, and finally understood the goal of the task. It had gone well since then."

P8 described how he adjusted his ways of doing task according to the timely feedback:

> "There was a time when the color suddenly turned red. This made me realize that my method might be wrong. Then I gradually adjusted my methods according to the variation trend of background color. After several attempts, I finally got the idea of the task. "

Apparently, the feedback, especially when indicating negative results, did make participants pause to think and study more.

4.3 Effects on Work Experience

The feedback of AuDi had even more impact on work experience. They reported how it provided them a way to evaluate their work on their own and adjust themselves accordingly, making them feel more in control and assured. At the same time, the change of color also easily evoked emotional responses from them, helping them to engage with the task or decide to quit eventually.

There is a big difference in participants' estimation of their own accuracy rate between the experiment group and the control group. That is, the control group's self-estimation was significantly lower: according to the questionnaire data, the average estimated accuracy rate given by the experiment group is 89.93%, while it is 79.56% by the control group. After subtracting estimated accuracy rate from the real accuracy rate, we got Figure 4. From this figure, we noticed that the estimated rate was more consistent with the real performance in the experiment group than the control group. As a matter of fact, there were several outliers in the control groups who unnecessarily considered their performance fairly poor: three between 0.41 and 0.6, and one even at 0.21, while all of these four actually performed far better. Our interviews also illustrated that with AuDi, the experiment group had much higher assurance of their performance than the control group. That is, with real-time feedback, AuDi eliminated their feelings of uncertainty or insecurity to a great extent. P10 shared her experience:

> "There were several questions that I found hard to judge. Without the feedback, I wouldn't be able to determine whether I made the right choice or not."

In addition, our participants reported that AuDi made them engaged with the task more. For instance, P1 described:

> "The feedback system helped me a lot to assess my work. I find it provides valuable information for my reference because it somehow interacts with my own thought and urges me to think about the standard to determine right or wrong."

More so, the auto-feedback display also evoked emotional responses, further urging or encouraging workers to find the right direction for doing the work. In our interviews, some reported that seeing low accuracy rate from the feedback display put pressure on them and evoked negative feelings such as anxiety or frustration, which, then, pushed them to work harder in trying and making the right decision so as to lift up the accuracy rates. One example was from PL:

> "The red color made me feel frustrated a little bit and urged me to make sure the next answer is right to change that situation."

However, interestingly, while they reported negative emotional responses when seeing negative feedback, they at the same time expressed positive feelings towards the feedback display. P8 put it this way:

> "When I saw it was red on my screen, I kind of felt relieved. Simply knowing that there was a mechanism detecting my potential errors made me feel secure and urged me to answer prudently. It's like only when you touch the 'bottom' line, can you learn to climb upwards easily."

On the other hand, if the color was always green, indicating fairly good work performance, they would feel more at ease and confident, as expressed by P3:

> "When the color was green or buff, I know I am good enough to get paid, and my worries and anxieties were gone and I would speed up prudently."

However, for those who couldn't find the right direction after several trials and errors, seeing redness all the time also pushed them to quit the project all together. For example, P11 who quit eventually reported:

> "The full-screen redness made my heart uncomfortable. In the following 5 problems, the screen didn't turn any greener. I felt that I couldn't raise the accuracy rate since the questions were totally beyond me. So I decided to quit."

Other participants also revealed that they would quit if they saw the red color all the time. Specifically, when the screen stayed red for several questions, indicating that the accuracy rate was below 80% , as such they thought that they would never understand the task requirement, let alone getting paid, so they were thinking of

quiting. After several tries, the color turned green again, so they decided to continue. This explains why the drop-out rate is slightly higher in the experiment group than that of control group, as they were sure they couldn't meet the requirement. As Table II shows, 3 participants of experiment group chose to quit, while only 1 in the control group dropped out at last.

People's emotional responses may also have something to do with the particular design of the feedback display. That is, the large area on the screen used to display color inevitably drew people's attention and somehow created a sense of immersion, which made them sensitive to the color change and the color itself. Besides, with the color schema of traffic lights, the related color did the right work to draw people the awareness and response, red for urgency and heightened awareness, green for relief and so on. That is, the color display provided participants with a rough but instant notion of the accuracy rate, leading to corresponding responses in a straightforward manner.

5 Discussions

As shown in our findings, AuDi, by automatically providing feedback in real time, helped engage our participants more and steer them to find the right direction for accomplishing the tasks, which then led to better performance in the end. At the same time, it also helped them feel more assured of the work. Overall, the quantitative and qualitative data analysis shows that the employment of AuDi was very well perceived by our participants, showing that it helped improve their work performance as well as work experiences in crowdsourcing.

Our findings of the use of AuDi indicates several advantages of the auto-feedback approach, compared to other automatic mechanisms, for crowd work performance and experience.

First, as shown in the data, although this auto-feedback approach does not explicitly ask or force people to pause and check, seeing the feedback itself, especially negative feedback, leads participants to actually pause, reread the task requirement, and put more thoughts and efforts to try to get things right. Compared to other intervention mechanisms such as inserting check stage (Wiseman et al., 2013), and introducing lockout (Gould et al., 2016), this auto-feedback mechanism provides more control and more autonomy to the hands of the workers, for them to decide on their own to take actions and do adjustments. As such, it was a more graceful, more humane, and more effective approach to engage workers to do the work right.

More so, the use of AuDi, while helping inform workers to make changes, does not introduce new interruptions, and as such largely protects the flow of work, very important for work performance. Studies show while some intervention mechanisms do improve work quality, the interventions shall be used cautiously as it may interrupt and disturb one's flow of work, and may instead have negative effects on work performance (Gould et al., 2016; Wiseman et al., 2013; Dai et al., 2015; Zhang et al., 2018). In general, workers might have difficulty in resuming to

perform tasks after experiencing an interruption and have to take time to regain focus (Iqbal and Horvitz, 2007) or suffer more stress and frustration in order to re-engage in less time (Mark et al., 2008). That is, with inappropriate interventions, interruption cost incurred in switching attention between tasks and as such would negatively affect workers' performance. When AuDi is considered, by simply providing feedback in the ambient form, without inserting breaks or lockout, it greatly minimizes the disturbing effects, and their flow of work protected.

Finally, the use of auto-feedback also appears to potentially relieve the commonly reported tensions between workers and requesters on crowdsourcing. Crowd workers are usually regarded as inexhaustible and anonymous labors, and were managed as such. The criteria of tasks are defined by the requesters and they have the final say in whether to accept the work and pay for them or not (Irani and Silberman, 2013). In consequence, workers are at risk of work rejection and have no reasonable resources to avoid this wage theft (McInnis et al., 2016; Irani and Silberman, 2013). Andrew Mao et al. investigated the reasons why workers drop out a crowd work and found out the most important reason is workers worried about their submitted answers being rejected (Mao et al., 2013). As such, rejecting their work without any reason or feedback was a commonly complained issue in crowdsourcing, and caused a lot of tensions between workers and requesters.

As shown in our study, the use of AuDi, by feeding the performance information back to the workers, not the requesters, quite successfully addressed workers' concern about quality. As reported by our participants, AuDi made our participants more aware of what they were doing in real time, and helped them make decisions on their own whether to go ahead confidently, pause to find ways to fix things, or to even quit completely. That is, what matters is not whether their work is rejected or not, but the reason of why the work is rejected, and AuDi is certainly helpful in that respect.

6 Limitations and Future Work

Though our study shows very positive results about using AuDi, there are also a number of limitations of the system and the study. First, the accuracy calculation method used makes it only work for those tasks with multiple choice questions, and not other tasks. So for crowd work that does not meet this requirement, AuDi can't apply without necessary adaption.

In addition, the particular algorithm might also lead to cold start effect. That is, AuDi's feedback is based on comparing submitted answers to the estimated right answers, so it relies on the already submitted answers to do the estimation. Owing to that, the first few workers will not get feedback as effective as the later ones do, as there are no other answers yet. Whether increasing transparency (e.g. displaying how many submitted answers on which the feedback is based) might be a good solution to this issues still needs further investigations.

Besides, there is also accumulated effect with the approach of AuDi. That is, the accuracy rate and the corresponding background color will change more dramatically at the beginning. As the number of questions answered grow, the accuracy rate will not be so greatly affected by one single answer anymore, so the color change will become less obvious. This effect was noticed and was also reported by our participants in the study, as they could see more background change at the beginning but not so much towards the end. To address this issue, we might divide all questions into multiple subgroups and to initialize the algorithm every time with each subgroup. But at the same time, people would rely more on the feedback at the beginning as learning is more actively taken by workers at the beginning. So it takes further investigations to find out whether the accumulated effect on the display shall be addressed and how.

7 Conclusions

In this paper, we presented an auto-feedback display called AuDi, as well as an experimental study to investigate how AuDi might work for crowdsourcing. Our study shows that people perceived the automatically calculated accuracy feedback as generally acceptable, and the feedback display itself was helpful for them to engage with the tasks and perform the work better. More specifically, it helped raise people's awareness and leading people to pause for thought and do the work more carefully when seeing red color, and encouraging them to proceed with more confidence when seeing green color. Without introducing new cost, AuDi shares the similar positive effects as personal feedback.

Hata et al.'s study shows that a worker's long-term performance is quite stable, as they usually adopt a particular strategy for completing tasks and will continue to use that strategy without change (Hata et al., 2017). However, as shown in our study, this is only true when there is no feedback for their work. When feedback is provided, as the use of AuDi in our case, changes of strategies for better performance could happen over the process.

*

References

Brewer, R., M. R. Morris, and A. M. Piper (2016): '"Why would anybody do this?": Understanding Older Adults' Motivations and Challenges in Crowd Work'. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. San Jose, California, USA, pp. 2246–2257.

Callison-Burch, C. (2009): 'Fast, cheap, and creative: evaluating translation quality using Amazon's Mechanical Turk'. In: *Proceedings of the 2009 Conference on Empirical Methods in Natural Language*. Singapore, pp. 286–295.

Dai, P., J. M. Rzeszotarski, P. Paritosh, and E. H. Chi (2015): 'And Now for Something Completely Different: Improving Crowdsourcing Workflows with Micro-Diversions'. In: *Proceedings of*

the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing. Vancouver, BC, Canada, pp. 628–638.

Dow, S., A. Kulkarni, S. Klemmer, and B. Hartmann (2012): 'Shepherding the crowd yields better work'. In: *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. Seattle, Washington, USA, pp. 1013–1022.

Feng, J., G. Li, H. Wang, and J. Feng (2014): 'Incremental Quality Inference in Crowdsourcing'. In: *International Conference on Database Systems for Advanced Applications*. pp. 453–467.

Franklin, M. J., D. Kossmann, T. Kraska, S. Ramesh, and S. Ramesh (2011): 'CrowdDB: answering queries with crowdsourcing'. In: *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*. Athens, Greece, pp. 61–72.

Gadiraju, U., R. Kawase, S. Dietze, and G. Demartini (2015): 'Understanding Malicious Behavior in Crowdsourcing Platforms: The Case of Online Surveys'. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Seoul, Republic of Korea, pp. 1631–1640.

Gould, S. J., A. L. Cox, D. P. Brumby, and A. Wickersham (2016): 'Now Check Your Input: Brief Task Lockouts Encourage Checking, Longer Lockouts Encourage Task Switching'. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. San Jose, California, USA, pp. 3311–3323.

Ham, J. and C. Midden (2010): 'Ambient persuasive technology needs little cognitive effort: the differential effects of cognitive load on lighting feedback versus factual feedback'. In: *Proceedings of the 5th international conference on Persuasive Technology*. Copenhagen, Denmark, pp. 132–142.

Hata, K., R. Krishna, L. Fei-Fei, and M. S. Bernstein (2017): 'A Glimpse Far into the Future: Understanding Long-term Crowd Worker Quality'. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. Portland, Oregon, USA, pp. 889–901.

Ipeirotis, P. G., F. Provost, and J. Wang (2010): 'Quality management on Amazon Mechanical Turk'. In: *Proceedings of the ACM SIGKDD Workshop on Human Computation*. Washington DC, pp. 64–67.

Iqbal, S. T. and E. Horvitz (2007): 'Disruption and recovery of computing tasks: field study, analysis, and directions'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. San Jose, California, USA, pp. 677–686.

Irani, L. C. and M. S. Silberman (2013): 'Turkopticon: interrupting worker invisibility in amazon mechanical turk'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Paris, France, pp. 611–620.

Janine, B. (2016): 'Income Security in the On-Demand Economy: Findings and Policy Lessons from a Survey of Crowdworkers'. *Comparative Labor Law & Policy Journal*, vol. 37, no. 3.

Kittur, A., J. V. Nickerson, M. Bernstein, E. Gerber, A. Shaw, J. Zimmerman, M. Lease, and J. Horton (2013): 'The future of crowd work'. In: *Proceedings of the 2013 conference on Computer supported cooperative work*. San Antonio, Texas, USA, pp. 1301–1318.

Kulkarni, A., M. Can, and B. Hartmann (2012): 'Collaboratively crowdsourcing workflows with turkomatic'. In: *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. Seattle, Washington, USA, pp. 1003–1012.

Little, G., L. B. Chilton, M. Goldman, and R. C. Miller (2010): 'TurKit: human computation algorithms on mechanical turk'. In: *Proceedings of the 23nd annual ACM symposium on User interface software and technology*. New York, New York, USA, pp. 57–66.

Liu, X., M. Lu, B. C. Ooi, Y. Shen, S. Wu, and M. Zhang (2012): 'CDAS: a crowdsourcing data analytics system'. *Proceedings of the VLDB Endowment*, vol. 5, pp. 1040–1051.

Mankoff, J., A. K. Dey, G. Hsieh, J. Kientz, S. Lederer, and M. Ames (2003): 'Heuristic Evaluation of Ambient Displays'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 169–176.

Mao, A., E. Kamar, Y. Chen, E. Horvitz, M. E. Schwamb, C. J. Lintott, and A. M. Smith (2013): 'Volunteering Versus Work for Pay: Incentives and Tradeoffs in Crowdsourcing'. In: *Proceedings of First AAAI Conference on Human Computation and Crowdsourcing*. Palm Springs, California, USA.

Mark, G., D. Gudith, and U. Klocke (2008): 'The cost of interrupted work: more speed and stress'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Florence, Italy, pp. 107–110.

Martin, D., B. V. Hanrahan, J. O'Neill, and N. Gupta (2014): 'Being a turker'. In: *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. Baltimore, Maryland, USA, pp. 224–235.

McInnis, B., D. Cosley, C. Nam, and G. Leshed (2016): 'Taking a HIT: Designing around Rejection, Mistrust, Risk, and Workers' Experiences in Amazon Mechanical Turk'. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. San Jose, California, USA, pp. 2271–2282.

Raykar, V. C., S. Yu, L. H. Zhao, G. H. Valadez, C. Florin, L. Bogoni, and L. Moy (2010): 'Learning from crowds'. *Journal of Machine Learning Research*, vol. 11, no. Apr, pp. 1297–1322.

Ross, J., L. Irani, M. S. Silberman, A. Zaldivar, and B. Tomlinson (2010): 'Who are the crowdworkers?: shifting demographics in mechanical turk'. In: *CHI '10 Extended Abstracts on Human Factors in Computing Systems*. Atlanta, Georgia, USA, pp. 2863–2872.

Rzeszotarski, J. M. and A. Kittur (2011): 'Instrumenting the crowd: using implicit behavioral measures to predict task performance'. In: *Proceedings of the 24th annual ACM symposium on User interface software and technology*. Santa Barbara, California, USA, pp. 13–22.

Wiseman, S., A. L. Cox, D. P. Brumby, S. J. Gould, and S. O'Carroll (2013): 'Using Checksums to Detect Number Entry Error'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Paris, France, pp. 2403–2406.

Zhang, Y., X. Ding, and N. Gu (2018): 'Understanding Fatigue and its Impact in Crowdsourcing'. In: *Proceedings of 2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design*. Nanjing, China, pp. 57–62.

Shekow, M.; Prinz, W. (2019): A capability analysis of groupware, cloud and desktop file systems for file synchronization. In: Proceedings of the 17th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing an the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep06

A capability analysis of groupware, cloud and desktop file systems for file synchronization

Marius Shekow and Wolfgang Prinz
Fraunhofer FIT, Sankt Augustin, Germany
Contact Author: marius.shekow|wolfgang.prinz@fit.fraunhofer.de

Abstract. Many groupware applications use hierarchical file systems, cloud storage or shared desktop operating system disks to support the cooperative development of shared artefacts or to share information. In these collaboration scenarios, often file synchronizers assist users in the data management across multiple devices. They establish consistency between file systems, even in light of their heterogeneity. However, the development of file synchronizers is difficult due to the fact that mainstream operating systems were not primarily built for cooperation or synchronization scenarios. Further, synchronizers need to address heterogeneity, by translating semantical differences and considering cross-device and cross-file system incompatibilities. This paper provides an in-depth analysis of six file system capabilities relevant to shared data synchronizers, such as mapping from namespace to physically stored objects, supported object types, namespace limitations or locking mechanisms. For each capability we derive commonalities for a set of selected file systems and also provide advice for handling incompatibilities. The insights of this work provide useful concepts and guidance for groupware developers that aim for a better user experience in synchronization support.

1 Introduction

With the increasing availability, affordability and mobility of computing devices like laptop computers, smartphones and tablets, working with multiple devices in

both professional and private life has become increasingly common. Users typically use applications like word processors or other domain-specific tools to create large parts of their data. The resulting documents are stored on the devices in a hierarchical *file system*, which has the role of a persistent database. Many different collaboration scenarios exist for files, such as users working together on office documents or file-based databases. To facilitate collaboration and to increase availability, documents are exchanged, e.g. via e-mail or via central systems, such as groupware, file servers or ubiquitous cloud storage. However, copying files and directories between storages causes problems, both for an individual user who wants to manage her files across these storage systems [1] and for collaborative multi-user scenarios.

One convenient solution is *data synchronization*, which has become popular, also due to the increased availability and affordability of cloud services (Yang et al., 2016). *File synchronizers*, as described by Balasubramaniam and Pierce (1998), are synchronizers whose data is the file system, including its namespace structure and file contents. In particular, cloud storage-based file synchronizers like Dropbox, Google Backup and Sync, OneDrive or NextCloud have become popular over the last ten years, indicated by the high number of their users (Kollmar, 2016; Price, 2017). They are programs that constantly run on a device in the background and tightly integrate with the file manager, providing a native user experience. They eliminate friction in file-based workflows because users no longer need to use 3rd party systems (such as a cloud storage web interface) but can work on the local file system directly, which avoids manual up- and downloads which cause files to lose their context (Vonrueden and Prinz, 2007). The offline availability of files improves navigation and search speed in the file system hierarchy. The native integration of a cloud synchronizer into the operating system and file manager provides many advantages. It makes 3rd-party functionality available at the user's finger-tips, such as the file manager's context menu which provides direct access to previous versions or comments of a file. Synchronizers also provide *synchronous awareness* (Fuchs et al., 1995), e.g. by showing native notifications in case new files were created, opened or locked by other users.

Today a plethora of industrial file synchronizers have emerged[2], used by a large user base. They need to support the file system APIs of all end-user operating systems they run on, as well as the API of the central file system. Building such file synchronizers is challenging, for several reasons. This work focuses on the fact that no two file systems are exactly equal, due to their heterogeneous capabilities. We use the term *capability* for a specific characteristic of a file system, such as namespace limitations or the way object relationships are modeled. Their traits may be different (heterogeneous) between any two file systems. If the

[1] Exemplary, users may fail to locate the correct, up to date version of a document on the right device. See e.g. Dearman and Pierce (2008); Jokela et al. (2015) for more details.

[2] E.g. Dropbox, Google Drive, Microsoft OneDrive, Amazon Drive, Box, NextCloud, Cloudstore, Resilio, Seafile, SpiderOakOne, LeitzCloud, Tonido, TeamDrive, MyDrive, Strato HiDrive, or Hubic. See (Wikipedia, 2017a) for a more complete list.

synchronizer developer ignores or overlooks a capability, this impairs the usability of the system because of bad side effects that occur during synchronization. Exemplary, if a developer overlooks that a file may not be named "aux" on Windows, the Windows implementation will run into unexpected loops or errors while trying to synchronize such a file, which was synchronized successfully by the macOS implementation. We have observed several instances of such side effects in practice in leading industrial synchronizers. The result is either "just" a divergence of the file systems, or worse, data loss.

We created this work as part of an ongoing endeavor to build a file synchronizer that overcomes the shortcomings of existing solutions supporting multi-user collaboration in asynchronous cooperation scenarios. We identify both homogeneous and heterogeneous capabilities relevant to file synchronizers. We propose suitable data transformation, where applicable, to avoid data loss. We start in section 2 where we briefly explain the mechanics of a file synchronizer and examine the variety of ways how file synchronizers define their file system. Next, we introduce five representative file systems we examined in section 3. In section 4 we present the detailed analysis of six capabilities. We conclude and present future work in section 5.

2 Background

Although the synchronization of information is essential for the support of collaborative work the CSCW research community focused primarily on researching synchronous synchronization and consistency algorithms such as the seminal work of Ellis and Gibbs (1989) on operation transformation and subsequent research by Sun and Ellis (1998); Sun and Sun (2009). On the other hand the CSCW community indicated the importance of consistent and contextual information sharing process (Voida et al., 2006). Although relevant for CSCW, file synchronization has primarily been researched in other domains. The seminal work by Balasubramaniam and Pierce (1998) describes and coins the term *file synchronizer*. We address authors of similar (or more powerful) synchronizers. A file synchronizer is a program that performs a pair-wise synchronization of two file system replicas upon the user's request, breaking synchronization down to a 3-stage process. In the first stage, update detection, the local and remote replicas are scanned to detect their current state. The list of changes (updates) is computed by comparing the current state to a locally persisted state from the point of the last synchronization. The second stage, reconciliation, is given the updates of both replicas and simulates (in memory) how the final, reconciled file system should look like which contains the updates of both replicas. The updates are examined for conflicts for which the user is asked to choose a suitable resolution. The output of this stage is the list of operations for the user to review in a graphical interface. The final stage, propagation, performs the actual file system modifications on each replica and updates the locally persisted state.

Such file synchronizers convert the file system from being a *collaboration-transparent*, replicated architecture to a *collaboration-aware* system (Phillips, 1999). Due to the heterogeneous capabilities of file systems, such synchronizers are also referred to as heterogeneous (Antkiewicz and Czarnecki, 2008; Foster et al., 2007). The advantage of heterogeneous synchronizers is that users can continue using existing file systems, without the (expensive) migration to a homogeneous system. The disadvantage is that the developer needs to build an internal model that is as compatible as possible with every file system the synchronizer aims to support. This involves finding a set of *common* capabilities, which we are doing in this work. Typically, the synchronizer *transforms* the heterogeneous model of each file system to the internal one. The transformation is challenging, because a suitable alignment needs to be found. The synchronizer then decides which updates to synchronize using the internal model. In extreme cases parts of the data are lost due to lack of alignment, as our work will show.

While there is a large number of industrial file synchronizers, the body of academic works is much smaller. We examined whether related works define a formal and thorough specification of their file system model, because we consider a formal definition of the data schema and its rules a basic requirement for any data synchronizer. Interestingly, a few works do not provide any specification of the file system and its operations, see e.g. (Cox and Josephson, 2005; Elijorde et al., 2013). Some provide a partial description, such as the record structure used to store the file system's state or the operations, see (Lindholm et al., 2005; Molli et al., 2003; Tao et al., 2015; Bao et al., 2011; Li et al., 2012; Uppoor et al., 2010). Others such as (Balasubramaniam and Pierce, 1998; Ng and Sun, 2016; Ramsey and Csirmaz, 2001; Csirmaz, 2016) formally specify a file system they *defined*. These works do not discuss the *mismatch* that exists between their internal model and the real-world file system their implementation actually works on.

Real-world file systems specifications, such as POSIX, are only formulated informally. A few academic works such as Ridge et al. (2015) exist which extracted exhaustive first-order logic (FOL) specifications for a few real-world implementations, but not all main stream file systems are covered yet. We present an *informal* comparison in this work instead, as this allows the provision of immediate results for a large selection of file systems. Some online resources such as (Craighead, 2008; Wikipedia, 2017b) also provide informal comparisons. Apart from (Jim et al., 2002), an unfinished manuscript by the authors of (Balasubramaniam and Pierce, 1998), there is no related scientific literature to the best of our knowledge that provides an in-depth discussion of the capabilities of file systems.

3 Examined file systems

To find capabilities we sample different *types* of file systems. As selection criteria we focus on market share and system type and chose one or two representative systems for each type. We examine Windows version 7-10 (NTFS) and macOS

version 10.11-10.13 (HFS+ and APFS) APIs because these are the most widespread end-user operating systems at the time of writing. Our findings also transfer to UNIX and therefore to both file servers (e.g. network-attached storage) and mobile devices such as smartphones. We consider WebDAV (Dusseault, 2007) which is widely available as interface for proprietary as well as open-source Internet (cloud) storages. Dropbox (HTTP API v2 (Dropbox Inc., 2017)) is chosen as a representative for widespread cloud storages (Dropbox Inc., 2016). *BSCW Social* (OrbiTeam Software GmbH & Co KG, 2018) is a representative for groupware systems commonly found in academia, a system that originates from the CSCW community (Bentley et al., 1997; Jeners and Prinz, 2014).

4 Capability analysis

This section provides an in-depth analysis of six capabilities relevant to file synchronizers. They were selected based on technical realities we discovered while implementing and technically evaluating a file synchronizer. Each capability is discussed in a separate subsection. For each one we first state its significance for the user, followed by an analysis, then extract similarities that manifest in the file synchronizer's internal model and finally give advice how file synchronizers can handle incompatibilities, if applicable.

4.1 Physical object & namespace mapping

The *namespace* is the user-facing side of a file system. It consists of a hierarchical set of *paths*, where a path is a notation for addressing a specific *object*. A path is a sequence of *names*, where names are simple strings. Hierarchy levels of a path are separated by a *separation character*, such as '/' or '\'. File system implementations differ in their approach how objects are identified, physically stored and how the mapping between namespace and objects works.

4.1.1 Significance

From the user's perspective the synchronizer translates a prefix of the synchronized namespace between the local disk and the remote storage, e.g. *'C:\SyncFolder'* to *'https://server.com/synced'*. Users expect that the local disk's and the server's namespace match exactly. However, due to technical limitations (analyzed below) this is not always possible. A synchronizer that is aware of incompatibilities should find a suitable way to inform the user about namespace mismatches (Dourish, 1996).

4.1.2 Analysis

An overview of the analysis is shown in figure 1.

We first classify whether file system objects (files, directories, etc.) can be identified uniquely (e.g. after moving them) by a persistent identity, or whether

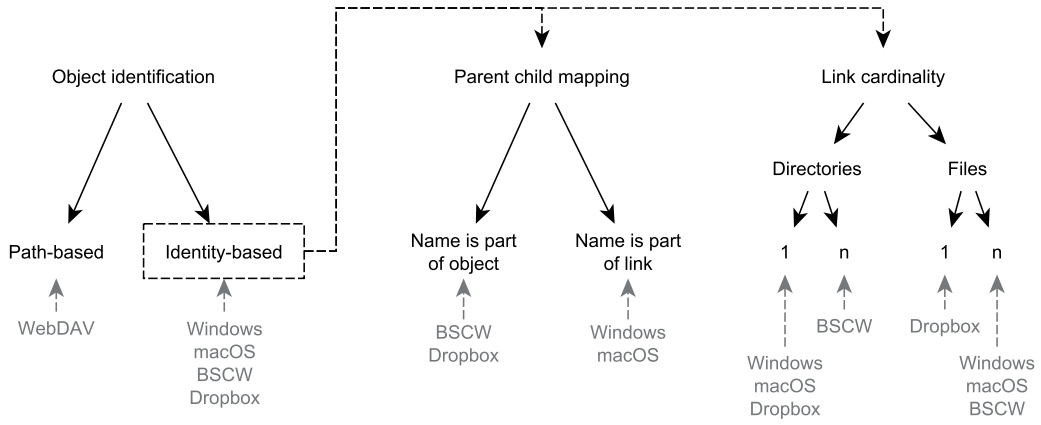


Figure 1. Analysis of object identification and namespace to object mappings.

only the path is available (Tao et al., 2015). Exemplary, Windows provides the *file index*, and macOS or UNIX systems provide *inode* numbers. For identity-based systems, two further classifications are appropriate, because an object with a specific ID may be accessible from one or more paths. In practice the cardinality varies per object type, s.t. Windows or macOS forbid more than one link to a directory to prevent cycles to occur in the tree. Some systems model the parent child relationship s.t. each directory has a list of $(name, id)$ tuples of its immediate children (name of the objects is part of the *link*), whereas others store the name as part of the object and each directory maintains a simple list of immediate child IDs.

Two more aspects not covered in figure 1 are that the invariants of each file system need further examination. A file system may or may not allow two sibling objects to have the same name, and it may use a case-sensitive or case-insensitive comparison while enforcing this invariant.

4.1.3 Derived unified model

To derive the internal file system model we suggest the following approach:

- If one or more file systems are *path-based*, either let the internal model be path-based too, or emulate IDs by generating them on the client, setting IDs as custom meta-data, if the file system API supports it (e.g. WebDAV PROPPATCH, see section 9.2 of Dusseault (2007)).
- When the parent child mapping varies, let the name be part of the object.
- If link cardinality varies, use the smaller (1) cardinality.
- When invariants vary, enforce the one that is most strict.

4.1.4 Advice for handling incompatibilities

When a file synchronizer encounters an incompatible mapping at run-time, e.g. if a specific file exists at multiple paths but the internal model limits the cardinality to 1, we suggest the synchronizer either stops synchronizing, asking the user to fix

the situation, or to automatically add the affected paths or IDs to an *ignore list*. Numerous industrial synchronizers provide such an ignore list users can fill with paths to files or directories they want to exclude from synchronization. We suggest that this list can also be manipulated by the reconciliation algorithm automatically to handle compatibility issues, notifying the user in such an event. For certain traits, workarounds may be possible. Exemplary, junctions (Windows) and symbolic links (macOS) may be used to allow a N-cardinality for directories. The synchronizer needs to choose one path as primary and use junctions or links for all other paths, updating them in case the primary path changes.

4.2 Supported object types

Files and directories are the two object types offered by all examined file systems. Jeners et al. (2013) show that even in groupware systems such as BSCW which offer many additional object types, the majority (90%) of user interaction takes place with these two object types. A file system may also support other object types that are incompatible with other systems.

4.2.1 Significance

When an object available on one file system is unavailable on the other one, its omission in the namespace, which is a loss of information, will confuse the user.

4.2.2 Analysis

While all examined file systems offer files and directories, there are several other types supported by just a subset of file systems, e.g. device files or symbolic links on macOS and Windows, or special types like contact lists, calendars or URLs on BSCW.

4.2.3 Derived unified model

By taking the intersection set of the available object types of each file system, the internal model should consist only of files and directories. We suggest to ignore other object types because they are specific to that file system and cannot be meaningfully viewed or manipulated on other systems that do not support them.

4.2.4 Advice for handling incompatibilities

We propose a similar handling as for mapping issues (section 4.1) where the synchronizer either stops or adds affected objects to the ignore list automatically, notifying the user about this action. A workaround is to create proxy objects, such as '.url' files, that allow the user to see the existence of the corresponding objects, redirecting the user to the respective location on the other file system in case she opens the proxy object.

4.3 Operations and atomicity

File system APIs offer many operations to both *query* the current state of the file system (e.g. listing a directory's content) or to *manipulate* it. In the update detection stage a file synchronizer relies on the query operations to extract the current state. At the final *propagation* stage, the synchronizer needs to transform the scheduled abstract operations (which equalize both file systems) to concrete operations of each file system. This is challenging because the exact operations, their preconditions and their degree of atomicity[3] vary.

4.3.1 Significance

A user expects that operations she applied to her local file system are consistently applied to other file systems by the synchronizer. Users also expect the synchronizer to avoid inconsistent states while synchronization is active or was interrupted. Not handling related issues causes confusion (e.g. attempting to open a partially transferred file) or additional work (such as manually cleaning up inconsistent files and directory structures) for the user.

4.3.2 Analysis

Every of the examined file systems offer operations to query the current state. The slight variations in query operation signatures are merely an implementation detail. When considering manipulation operations, all file systems offer operations to create or delete empty directories, or to move an object. However, there is significant variation in the availability and atomicity of operations used to create or update files, or to delete non-empty directories. Exemplary, BSCW allows to atomically create non-empty files or delete non-empty directories, while Windows does not. Another observation is that desktop file systems like Windows and macOS offer *mount* operations which create a mount point that establishes a transition between volumes.

4.3.3 Derived unified model

A user would expect a file synchronizer to be capable of a set of operations the user also knows from using the file manager. An exemplary list could be as follows:

- *createdir(path)* creates an empty directory at *path*
- *deletefile(path)* deletes the file at *path*
- *deletedir(path)* deletes the directory and all its children at *path*
- *move(source, dest)* moves an existing object from *source* to *dest*
- *transfer(source, dest)* transmits a *file* located at *source* on the source file system to *dest* on the destination file system, to create a new file or update an existing one

[3] We refer to *atomicity* as known from database systems, see also section 1.3.4 of Elmasri and Navathe (2015).

To not leave either file system in an inconsistent state, every operation is expected to succeed or fail atomically. Optionally, a *copy file* operation can be used to copy a file on the destination file system in case it is feasible to detect exact copies of files on the source file system, e.g. by using checksums.

4.3.4 Advice for handling incompatibilities

All discrepancies we found between concrete file system operations and the ones presented above result from varying degrees of atomicity, which can be solved in the following ways:

- *deletedir(path)*: if a file system does not offer an atomic, recursive implementation, we suggest to first call *move(path, temp)* where *temp* is a path outside of the synchronized namespace, but on the same volume. This move operation succeeds (or fails) atomically and *appears* as an atomic delete operation to the synchronizer. Next, perform a *post-order* traversal of *temp*'s sub-namespace, deleting first files then directories.
- *transfer(source, dest)*: if the destination file system's operation is not atomic, we propose to execute *transfer(source, temp)*, i.e., write transferred data to a temporary location *temp* that is outside the synchronized namespace but also on the same volume. Once finished, perform *move(temp, dest)* on the destination file system.

Finally, file synchronizers which detect move operations via the object's ID should be aware of *mount points* within the synchronized namespace. IDs are only unique within a volume. However, a mount point establishes a transition between volumes. When the user performs a conceptual *move(source, dest)* operation where *source* is on volume *A* and *dest* on volume *B*, the synchronizer will incorrectly detect a *delete* operation for *source* and a *create* operation for *dest*. We therefore suggest that synchronizers detect mount points and either reject them (by stopping synchronization) or automatically adding them to the *ignore list*.

4.4 Namespace limitations

Although the general namespace consists of *Unicode* characters, a file system may pose limitations on the namespace, affecting paths or the names of a path, usually for technical or historical reasons.

4.4.1 Significance

When a user attempts to create an object with a name that violates a namespace limitation, the file manager (or web interface) prevents the creation and provides *immediate* feedback how to fix the name. When using file synchronization, the chosen name may be accepted by the source file system API, but may violate a limitation of the destination API. The file synchronizer discovers this issue after a (possibly large) delay which surprises the user, because to her the creation of the object initially appeared to be successful. Furthermore, users will be confused if

objects exist on one system but not the other one due to a limitation that affects only the latter system.

4.4.2 Analysis

The following list provides a brief summary of our findings. We refer to the respective file system documentation for further details[4].

- A file system may reserve a set of *characters* from being used in object names, either at any position, or only in specific positions. *Forward slashes* are forbidden in all examined systems, as they separate names in a path. Windows reserves the most characters, and other systems such as BSCW or Dropbox have adopted Windows' set of reserved characters and names for compatibility reasons.
- Similarly, some systems reserve a set of *names*, such as "." or "..". Windows reserves a large set of names such as "CON" or "PRN" for historical reasons and also reserves *short file names* (Microsoft Inc., 2018) in case a longer file name already exists (exemplary, given a directory named "project report", creating an object at "projec~1" is forbidden on volumes with short file name creation enabled).
- Many systems impose a maximum length of names and paths. Often names are limited to a length of 255 characters. Shorter path lengths (such as macOS with 1016 characters) also cause issues, e.g. deep directory hierarchies being in accessible.
- While all examined systems use the Unicode alphabet with some form of encoding (e.g. UTF-8), not all systems *preserve* the *normalization form* (such as NFC or NFD[5]) of characters. Exemplary, the HFS+ file system on macOS does not preserve a large set of input characters but converts them to a NFD-like form.
- *Case-sensitivity* may vary between two file systems. By default, the *examined* systems are all case-insensitive. However, others such as the UNIX file system, are case-sensitive! We found all systems to be case-*preserving*.
- In rare instances the file system APIs behave deceptively. They accept a name, seemingly execute successfully, but actually change the name internally. This is problematic for file synchronizers, as the next update detection phase will find an unexpected name and assume that the object was moved by the user. One example is the Unicode normalization conversion of HFS+ volumes mentioned above, another is Windows which silently strips trailing spaces/dots from a name during execution.

[4] See e.g. Berners-Lee et al. (1994), Apple Inc. (2004), Apple Inc. (2017) or Microsoft Inc. (2018).

[5] See *http://unicode.org/reports/tr15/*, retrieved January 2, 2019.

4.4.3 Derived unified model

For each limitation of file systems A and B we propose to take the one that is more strict and let the file synchronizer apply it to the file system with the weaker limitation. For reserved characters or names this means to apply the *union* of the sets to both A and B. For length limitations, the shorter length is more strict. Also, case-insensitivity is more strict than case-sensitivity.

4.4.4 Advice for handling incompatibilities

We suggest a file synchronizer takes one of the following approaches when encountering paths that are incompatible w.r.t. the unified limitations:

1. Stop synchronization, ask the user to manually rename objects
2. Automatically rename objects to establish compatibility
3. Automatically add incompatible objects to the ignore list

While approach (1) is easy to implement, it is labor-intensive for the user. In case the stopped synchronization goes unnoticed, and if it remains in that state for extended periods of time, this increases the chance for conflicts. Approach (2) mitigates this problem, but automatic renaming can cause issues when the affected objects belong to a naming scheme of a third party application. Such applications may stop working once these files and directories no longer correspond to the expected naming scheme. The last approach fixes the issues of the two ones but requires the implementation of the aforementioned ignore list.

4.5 Meta-data

Meta-data provides further information about objects. It is not stored as part of the object, but at a separate location.

4.5.1 Significance

When meta-data stored on one file system is incompatible with the other file system, a synchronizer must skip their synchronization or perform a conversion. This type of data loss negatively affects the user, because she cannot access meta-data available only on the remote file system during an offline period.

4.5.2 Analysis

Each file system provides a diverse set of meta-data. Some meta-data are *attributes* managed by the file system, others can be changed by a client applications, such as a file synchronizer. Some systems offer one or more APIs to write *custom* meta-data, e.g. *Extended Attributes* and *Alternate Data Streams* on Windows, or *xattr* and *Resource forks* on macOS. The following meta-data is available on *all* file systems:

- Object type (file, directory, ...)
- File size (for files)
- Timestamp of creation and last modification

4.5.3 Derived unified model

All file systems support the retrieval of meta-data that is necessary to extract their state, such as the object's type or the last-modified timestamp. In case a file synchronizer models the file system using IDs, all file systems except for WebDAV automatically generate and provide unique IDs. For WebDAV we propose that the *file synchronizer* generates globally unique IDs (GUIDs) when creating objects on a WebDAV file system, assigning the GUID via the *PROPPATCH* command.

4.5.4 Advice for handling incompatibilities

Some meta-data, such as *attributes*, are system-specific and often lose meaning when copied to another file system, especially when it is of different type or located on a different operating system or machine. Exemplary, synchronizing the *compressed* attribute of a Windows file to the corresponding file on a macOS file system defies any purpose. We find that bypassing meta-data synchronization largely facilitates a file synchronizer's implementation. This also applies to *authorization* mechanisms, such as UNIX *permissions* or the more powerful *Access Control List* entries, which can also be considered to be meta-data, with varying availability and heterogeneity.[6]

The *last-modified* timestamp is an exception. We suggest to synchronize it because it is typically available on each file system, has the same meaning everywhere and users are aware of it when using the file manager. A caveat developers need to consider is the variety of resolutions and formats of timestamps.

4.6 Locking

Locking allows one user to *exclusively* modify an object on a file system, while all other users are prevented from modifying their own replica of that object.

4.6.1 Significance

Locking is an important mechanism that introduces *pessimistic* concurrency control in situations where users expect that conflicts are likely to happen. It avoids conflicts or lost updates. In an example scenario, a user locks a document she exclusively wants to work on for an hour. During this time, other users should be unable to concurrently modify this file, and should be *aware* of this lock while it is set. The

[6] As an example for heterogeneity, macOS and Windows both support *Access Control Lists*, but their implementations vary considerably. Additionally, synchronization of authorization data would require to also synchronize *authentication* data, i.e., user accounts, which introduces additional challenges.

information about the lock's existence can be propagated by the synchronizer to other users while they are online. In practice we have not observed locking to play a role for files stored on *local* disks. However, this feature is frequently used in groupware systems such as BSCW, and the transparent handling and awareness of locking behavior is an early requirement for CSCW systems as described in Blair and Rodden (1994).

4.6.2 Analysis

We analyzed the file systems' locking capabilities to determine whether a file synchronizer can safely protect an object from modification by the local user, because a different user locked the object. We found that some systems such as Dropbox do not offer any locking mechanism. Systems such as WebDAV and BSCW provide an elaborate locking model, including lock meta-data such as the owner and expiration time.

The locking mechanisms of Windows (*read-only* attribute, *file handle* locking) and macOS (*immutable* attribute, advisory locks via *fcntl*[7] API) are less elaborate. They each work differently and protect other aspects of modification. Exemplary, the read-only attribute on Windows does not protect objects from being moved or renamed, while the immutable attribute on macOS does.

We think that this diversity stems from the fact that each mechanism has a different purpose. On Windows and macOS the read-only/immutable file attribute or handle-based locks were not designed for a multi-user locking scenario. It is our understanding that they exist to allow users (and programs) to protect objects from modification *on the same device*, not across multiple devices. Handle-based locking suffers from *volatile* characteristics[8]. On macOS, handle-based locking is designed for a set of *cooperating* programs and not intended to prevent third party programs from modifying files. On Windows, handle-based locking has more wide-spread effect than just locking the object itself. It works on a "first come, first served" basis. Even just opening a file for reading already locks it. A file synchronizer may fail to obtain a lock, or inadvertently lock the path of any parent object, which is not desired. In addition, reliable recursive locking of a *directory* is not possible with the mechanisms offered by Windows and macOS.

4.6.3 Derived unified model

In case a pair-wise synchronization targets two file systems of *equal* type, such as two WebDAV systems, lock synchronization is feasible. In any other scenario we advise to ignore lock synchronization due to the strong differences in their implementation, making it impossible to meaningfully map one lock type onto another one.

---

[7] *http://man7.org/linux/man-pages/man2/fcntl.2.html*, retrieved January 2, 2019.

[8] When the program that owns the handle to an object terminates, the lock is automatically cleared.

4.6.4 Advice for handling incompatibilities

Not synchronizing locks does not necessarily mean that the synchronizer completely ignores locks. Some systems like WebDAV allow the *discovery* of locks (*before* the attempt of modifying a locked resource). Assume a scenario where a synchronizer detects that the user updated file f locally, while f is locked on the remote replica by another user. The synchronizer may then skip synchronizing f and notify the user about the lock's existence. With additional implementation effort, a synchronizer may also monitor the user's opened files and warn her in case she opens a file that is locked by other users. It is also possible to convey the existence of locks by the use of *overlay icons* in the file manager.

If lock discovery is unavailable we propose to treat failures like any other permission-related ones, such as failures resulting from prohibitive ACL entries or UNIX permissions. The synchronization may be stopped or the affected object could be skipped. The user should be notified about the problem in either case and be provided with as much available information as possible to fix the problem.

4.7 Summary

A summary of the capabilities of each file system is shown in figure 2. This radar chart depicts a rough estimate of the degree of power for each capability from 0% (center) to 100%, based on a technical evaluation beyond the scope of this work. Smaller values indicate less powerful namespace mappings, fewer supported object types, stronger namespace limitations, smaller level of locking, etc. We chose 20% as minimum value only to improve readability. By intersecting the areas of the file systems a synchronizer supports we can derive the degree of limitations of the synchronizer's internal model.

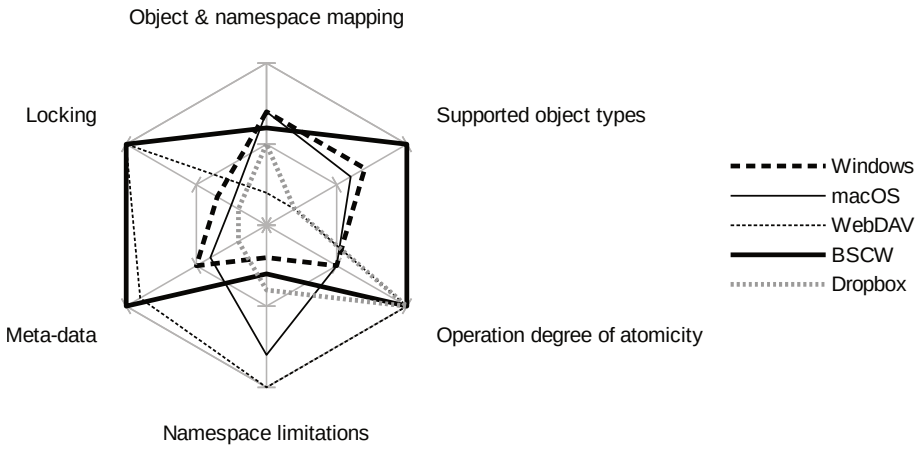


Figure 2. File system capabilities overview.

5 Conclusions and future work

In this work we have analyzed several file system capabilities relevant to file synchronizers as a baseline for the development of cooperation support applications. Synchronizers facilitate data management and collaboration in single- and multi-user settings. Supporting a variety of heterogeneous systems satisfies the user's need to synchronize between different devices and services, thus aiming at the provision of an integrated collaboration environment (Prinz et al., 2009).

Synchronization of the examined file systems is challenging due to their heterogeneity. This first and foremost affects the structure of a file system. As we discussed in section 4.1 and 4.2 two file systems may vary how paths of the namespace are mapped to objects or which types of objects exist. We proposed to generally take the lowest common denominator, e.g. limit synchronization to files and directories, or allow each object to be linked just once into the namespace. We proposed that the most user-friendly solution to deal with incompatible paths is to add them to an *ignore list* automatically. In subsection 4.3 we informally presented a set of commonly available operations that are sufficient to achieve consistency. Some of them require a degree of *atomicity* not offered by some implementations like Windows and macOS. For these we provided workarounds which emulate atomic behavior. In section 4.4 we found that Windows is imposing strong namespace limitations due to a large set of reserved names and characters. BSCW and Dropbox mimic Windows' behavior for compatibility reasons. Consequently, objects with incompatible names need to be dealt with, for which we presented several approaches, each with their own advantages and disadvantages. We analyzed accessible meta-data in section 4.5. Except for WebDAV, all file systems provide an automatically generated object ID which allows a file synchronizer to uniquely identify objects irrespective of their path, which facilitates the detection of move operations. Except for the *last-modified* timestamp we consider synchronization of other meta-data inadequate. Finally, section 4.6 discusses locking. We find that lock semantics of two file systems of different type are too heterogeneous to allow for a meaningful lock synchronization. Where possible, synchronizers should provide *awareness* of active locks to the user.

Despite the discussed caveats we still consider the use of file synchronization an enrichment of the user's experience. We hope that authors and developers of file synchronizers find our in-depth analysis and advice useful when implementing heterogeneous file synchronizers. While a lot of the given advice for handling incompatibilities may appear straightforward, our analysis of several industrial file synchronizers has shown a great variety in behavior, including many illogical choices[9]. We also hope that developers of next-generation file systems may also find clues to build systems that better support synchronization, in particular considering aspects such as *locking* in a cooperative setting. As future work we

[9] Exemplary, the macOS implementation of OneDrive uploads files with Windows-reserved *names* without warning, but skips synchronization of reserved *characters*.

will analyze the effects of how a file synchronizer models the state and operations of a file system on the conflicts that it detects, including a discussion of conflict resolution approaches taken by different related works. A user study is planned to verify our recommendations of handling incompatibilities with users.

References

Antkiewicz, M. and K. Czarnecki (2008): 'Design Space of Heterogeneous Synchronization'. In: R. Lämmel, J. Visser, and J. Saraiva (eds.): *Generative and Transformational Techniques in Software Engineering II: International Summer School, GTTSE 2007, Braga, Portugal, July 2-7, 2007. Revised Papers*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 3–46.

Apple Inc. (2004): 'Technical Note TN1150: HFS Plus Volume Format'. https://developer.apple.com/legacy/library/technotes/tn/tn1150.html.

Apple Inc. (2017): 'Apple File System Guide - FAQs'. https://developer.apple.com/library/content/documentation/FileManagement/Conceptual/APFS_Guide/FAQ/FAQ.html.

Balasubramaniam, S. and B. C. Pierce (1998): 'What is a File Synchronizer?'. In: *Proceedings of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking*. New York, NY, USA, pp. 98–108, ACM.

Bao, X., N. Xiao, W. Shi, F. Liu, H. Mao, and H. Zhang (eds.) (2011): 'SyncViews: Toward Consistent User Views in Cloud-Based File Synchronization Services: 2011 Sixth Annual Chinagrid Conference'.

Bentley, R., T. Horstmann, and J. Trevor (1997): 'The World Wide Web as Enabling Technology for CSCW: The Case of BSCW'. *Computer Supported Cooperative Work (CSCW)*, vol. 6, no. 2, pp. 111–134.

Berners-Lee, T., L. Masinter, and M. McCahill (1994): 'Uniform Resource Locators (URL)'.

Blair, G. S. and T. Rodden (1994): 'The Challenges of CSCW for Open Distributed Processing'. In: *Proceedings of the IFIP TC6/WG6.1 International Conference on Open Distributed Processing II*. Amsterdam, The Netherlands, The Netherlands, pp. 127–140, North-Holland Publishing Co.

Cox, R. and W. Josephson (2005): 'File Synchronization with Vector Time Pairs'.

Craighead, M. (2008): 'Windows vs. Unix File System Semantics'. http://www.conifersystems.com/2008/10/21/windows-vs-unix-file-system-semantics/.

Csirmaz, E. (2016): 'Algebraic File Synchronization: Adequacy and Completeness'. https://arxiv.org/pdf/1601.01736.pdf.

Dearman, D. and J. S. Pierce (2008): 'It's on my other computer! Computing with multiple devices'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Florence, Italy: ACM, pp. 767–776.

Dourish, J. P. (1996): 'Open implementation and flexibility in CSCW toolkits'. phd, University of London, London.

Dropbox Inc. (2016): 'Celebrating half a billion users'. https://blogs.dropbox.com/dropbox/2016/03/500-million/.

Dropbox Inc. (2017): 'Dropbox for HTTP Developers: The Dropbox API v2'. https://www.dropbox.com/developers/documentation/http/overview.

Dusseault, L. (2007): 'HTTP Extensions for Web Distributed Authoring and Versioning (WebDAV)'.

Elijorde, F. I., H. Yang, and J. Lee (2013): 'Ubiquitous Workspace Synchronization in a Cloud-based Framework'. *Journal of Korean Society for Internet Information*, vol. 14, no. 1, pp. 53–62.

Ellis, C. A. and S. J. Gibbs (1989): 'Concurrency control in groupware systems'. In: *Proceedings of the 1989 ACM SIGMOD international conference on Management of data*. Portland, Oregon, USA: ACM, pp. 399–407.

Elmasri, R. and S. B. Navathe (2015): *Fundamentals of database systems*. Pearson.

Foster, J. N., M. B. Greenwald, C. Kirkegaard, B. C. Pierce, and A. Schmitt (2007): 'Exploiting schemas in data synchronization'. *Journal of Computer and System Sciences*, vol. 73, no. 4, pp. 669–689.

Fuchs, L., U. Pankoke-Babatz, and W. Prinz (1995): 'Supporting Cooperative Awareness with Local Event Mechanisms: The GroupDesk System'. In: H. Marmolin, Y. Sundblad, and K. Schmidt (eds.): *Proceedings of the Fourth European Conference on Computer-Supported Cooperative Work ECSCW '95: 10–14 September, 1995, Stockholm, Sweden*. Dordrecht: Springer Netherlands, pp. 247–262.

Jeners, N., O. Lobunets, and W. Prinz (2013): 'What groupware functionality do users really use? A study of collaboration within digital ecosystems'. In: *2013 7th IEEE International Conference on Digital Ecosystems and Technologies (DEST)*. pp. 49–54.

Jeners, N. and W. Prinz (2014): 'Metrics for Cooperative Systems'. In: *Proceedings of the 18th International Conference on Supporting Group Work*. Sanibel Island, Florida, USA: ACM, pp. 91–99.

Jim, T., B. C. Pierce, and J. Vouillon (2002): 'How to build a file synchronizer'. https://www.cis.upenn.edu/ bcpierce/courses/dd/papers/unisonimpl.ps.

Jokela, T., J. Ojala, and T. Olsson (2015): 'A Diary Study on Combining Multiple Information Devices in Everyday Activities and Tasks'. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Seoul, Republic of Korea: ACM, pp. 3903–3912.

Kollmar, F. (2016): 'The Cloud Storage Report – Dropbox Owns Cloud Storage on Mobile'. https://blog.cloudrail.com/cloud-storage-report-dropbox-owns-cloud-storage-mobile/.

Li, Q., L. Zhu, S. Zeng, and W. Q. Shang (eds.) (2012): 'An Improved File System Synchronous Algorithm: 2012 Eighth International Conference on Computational Intelligence and Security'.

Lindholm, T., J. Kangasharju, and S. Tarkoma (2005): 'A hybrid approach to optimistic file system directory tree synchronization'. In: V. Kumar, A. Zaslavsky, U. Cetintemel, and A. Labrinidis (eds.): *The 4th ACM international workshop on Data engineering for wireless and mobile access*. New York, NY, USA, pp. 49–56, ACM.

Microsoft Inc. (2018): 'MSDN documentation: Naming Files, Paths, and Namespaces'. https://msdn.microsoft.com/en-us/library/windows/desktop/aa365247(v=vs.85).aspx.

Molli, P., G. Oster, H. Skaf-Molli, and A. Imine (2003): 'Using the transformational approach to build a safe and generic data synchronizer'. In: *Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work*. Sanibel Island, Florida, USA: ACM, pp. 212–220.

Ng, A. and C. Sun (2016): 'Operational Transformation for Real-time Synchronization of Shared Workspace in Cloud Storage'. In: *Proceedings of the 19th International Conference on Supporting Group Work*. Sanibel Island, Florida, USA: ACM, pp. 61–70.

OrbiTeam Software GmbH & Co KG (2018): 'BSCW Social'. https://www.bscw.de/social/.

Phillips, W. G. (1999): 'Architectures for synchronous groupware: Technical Report 1999-425'.

Price, R. (2017): 'Google Drive now hosts more than 2 trillion files'. http://www.businessinsider.de/2-trillion-files-google-drive-exec-prabhakar-raghavan-2017-5.

Prinz, W., N. Jeners, R. Ruland, and M. Villa (2009): 'Supporting the Change of Cooperation Patterns by Integrated Collaboration Tools'. In: L. M. Camarinha-Matos, I. Paraskakis, and H. Afsarmanesh (eds.): *Leveraging Knowledge for Innovation in Collaborative Networks*. Berlin, Heidelberg, pp. 651–658, Springer Berlin Heidelberg.

Ramsey, N. and E. Csirmaz (2001): 'An algebraic approach to file synchronization'. In: A. M. Tjoa and V. Gruhn (eds.): *the 8th European software engineering conference held jointly with 9th ACM SIGSOFT international symposium*. p. 175.

Ridge, T., D. Sheets, T. Tuerk, A. Giugliano, A. Madhavapeddy, and P. Sewell (2015): 'SibylFS: Formal specification and oracle-based testing for POSIX and real-world file systems'. In: *Proceedings of the 25th Symposium on Operating Systems Principles*. Monterey, California: ACM, pp. 38–53.

Sun, C. and C. Ellis (1998): 'Operational transformation in real-time group editors: Issues, algorithms, and achievements'. In: *Proceedings of the 1998 ACM conference on Computer supported cooperative work*. Seattle, Washington, USA: ACM, pp. 59–68.

Sun, D. and C. Sun (2009): 'Context-Based Operational Transformation in Distributed Collaborative Editing Systems'. *IEEE Transactions on Parallel and Distributed Systems*, vol. 20, no. 10, pp. 1454–1470.

Tao, V., M. Shapiro, and V. Rancurel (2015): 'Merging Semantics for Conflict Updates in Geo-distributed File Systems'. In: *Proceedings of the 8th ACM International Systems and Storage Conference*. New York, NY, USA, pp. 10:1–10:12, ACM.

Uppoor, S., M. D. Flouris, and A. Bilas (2010): 'Cloud-based synchronization of distributed file system hierarchies'. In: *2010 IEEE International Conference On Cluster Computing Workshops and Posters (CLUSTER WORKSHOPS)*. pp. 1–4.

Voida, S., W. K. Edwards, M. W. Newman, R. E. Grinter, and N. Ducheneaut (2006): 'Share and Share Alike: Exploring the User Interface Affordances of File Sharing'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 221–230, ACM.

Vonrueden, M. and W. Prinz (2007): 'Distributed Document Contexts in Cooperation Systems'. In: B. Kokinov, D. C. Richardson, T. R. Roth-Berghofer, and L. Vieu (eds.): *Modeling and Using Context*. Berlin, Heidelberg, pp. 507–516, Springer Berlin Heidelberg.

Wikipedia (2017a): 'Comparison of file synchronization software'.

Wikipedia (2017b): 'Comparison of file systems'.

Yang, C.-T., W.-C. Shih, C.-L. Huang, F.-C. Jiang, and W. C.-C. Chu (2016): 'On construction of a distributed data storage system in cloud'. *Computing*, vol. 98, no. 1, pp. 93–118.

Economidou Eleni, Krischkowsky Alina, Leitner Bianca, Murer Martin, Tscheligi Manfred (2019): On Middle-Ground Solutions for Domain-Specific Problems: The Case of a Data Transfer System for Sign Language Teachers. In: Proceedings of the 17th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing an the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep15

On Middle-Ground Solutions for Domain-Specific Problems: The Case of a Data Transfer System for Sign Language Teachers

Eleni Economidou, Alina Krischkowsky, Bianca Leitner,
Martin Murer, Manfred Tscheligi
Center for Human-Computer Interaction, University of Salzburg, Austria
firstname.lastname@sbg.ac.at

Abstract. Oftentimes domain-specific problems are imperceptible. The end users are so accustomed to the conditions that they do not request any other solution. In the context of a school for both hearing-impaired and hearing children, the particular way of teaching sign language led to the emergence of technology-mediated yet ill-supported work practices. This paper contributes to the CSCW community by introducing an approach for addressing domain-specific problems by applying a *middle-ground* yet optimal solution. As a direct outcome of this approach, we present the case of a data transfer system that supports sign language teachers' work practices. This system, which is indefinitely deployed in the school, is a tangible representation of current infrastructural and contextual issues the teachers are facing, and serves both as a reflection on the work practices and an articulation means of the limitations that constraint them. We reflect on our approach, we discuss on the resulted case in terms of an indefinitely deployed research product, and we speculate on the system's alternative application domains.

Keywords: domain specific problems, middle-ground solution, integrated school, work practices, sign language teachers, data transfer system.

1 Introduction

The starting point of this paper is the field research we conducted at an integration school in Salzburg[1], to gather a deeper understanding of the contextual particulars as well as the technological needs of the students and teachers. At this school, teachers and students with varying degrees of hearing abilities (i.e., hearing, hearing impaired, and deaf) are teaching and being taught. The uniqueness of this constellation is reflected in the school's exceptional way of teaching, educational philosophy, and subject matter. Beyond verbal communication, sign language is a key communication means among students and school staff. Consequently, at this school, sign language[2] is a compulsory subject, taught by two sign language teachers. Sign languages are fully-fledged natural languages with their own grammar and lexicon that make use of visual perception to convey meaning (Wilcox and Occhino, 2016). These particularities require teachers to develop new, innovative teaching material in terms of format, media, and content.

There is a substantial amount of work in both HCI and CSCW research with hearing-impaired and deaf communities and, in particular, with children (e.g., Slegers et al, 2010; Vermeulen et al, 2012). What unites these research efforts is their strong dedication to user-centred and participatory approaches to user involvement and design (e.g., Morningstar et al, 2015; Slegers et al, 2010; Vermeulen et al, 2012). In line with research suggesting strong user involvement, our presented research is also characterised by strong user participation and involvement; conducting participatory observations in classes or conducting semi-structured interviews with teachers. Throughout the conducted fieldwork, we identified that videos are considered as essential media and educational material to teach sign language, i.e., video footage of teachers to assign homework, or video footage of students to perform and document exercises, tests, or homework.

Currently, video file transfers are accomplished via an exchange of USB drives. In our research, we observed that the teachers' current practice of exchanging video files between them and the students results in a lot of effort since the copying process is both a tedious and time-consuming task. Other, more elaborated and state-of-the-art technology is already available on the market which supports this kind of data transfer (e.g., cloud or server-based solutions, or even a tailored local file transmission platform) allowing for an *ideal*[3] (Table I), yet not optimal, solution

---

1 The Josef Rehrl school comprises of an elementary and a secondary education school. https://www.josef-rehrl-schule.salzburg.at

2 Having sign language as a compulsory subject is exceptional and unique since Austrian sign language (ÖGS) was legally recognised as an official language by the Austrian Parliament on September 2005, and certified ÖGS translators are no more than 105. https://www.josef-rehrl-schule.salzburg.at/2014-05-12-09-52-56/methodik-und-leistungsziele.html

3 In this paper we make use of the adapted definition of the adjective *ideal* from the Oxford Learner's Dictionary of Academic English: "the best that can be imagined, but not likely to become real". https://www.oxfordlearnersdictionaries.com/definition/academic/ideal1, accessed February 1, 2019.

for the given problem space. The problematic case here is the school's current technological infrastructure, the lack of technical maintenance, and a shortage in teaching staff i.e., a lack of certified sign language teachers in schools, plus a lack of a community of practice to communicate and cooperate with. All these constraints provide the ÖGS teachers with a certain boundary box, forcing them to invent their own teaching practices and materials.

In our empirical research (i.e., participatory observations and semi-structured interviews) we identified this tension of an evolving new subject; while the technological infrastructure in the school, as well as the community-based exchange among ÖGS teachers is Austria, lags behind. In this paper we present a technological solution that from a research perspective may be perceived as the *middle-ground*[4] solution (Table I). Our *middle-ground* solution embodies both, the teachers' need of technical support in their data transfer practices, while being at the same time an *optimal* solution in the given context that colludes with the existing technological infrastructure in the school, rather than radically disrupts the infrastructure or the teachers existing work practices. It is important to mention that what is perceived as ideal or middle ground is very context and user specific. Therefore, throughout this paper, we indicate for whom the solution is perceived as ideal or middle ground. This technological solution is currently deployed at the school and used by the teachers.

Table I. In the design of work practice infrastructures striving for the ideal might not result in feasible solutions.

The ideal	**The *middle-ground***
An ideal solution is considered to be the theoretically perfect yet often unattainable opportunity.	A *middle-ground* solution is an attainable compromise between the ideal solution and the current boundary conditions.

The main goal of this paper is to articulate why addressing certain domain specific problems requires to also advance systems, that at first glance might seem outdated or superseded by the state of the art.

The contributions of this paper are threefold: 1) we present an approach for addressing domain-specific problems with *middle-ground* yet *optimal* solutions, 2) we present the case of the *design* of a fully functional system as a direct outcome of that process, and 3) we *reflect* on our pursued process and the developed system from diverse angles.

We first present related work to position our research in the realm of HCI and CSCW, we describe the process we pursued throughout our research, we continue

4 In this paper we make use of the adapted definition of the noun *middle ground* from the Collins English Dictionary: "a position of compromise between two opposing views, parties, etc". https://www.collinsdictionary.com/dictionary/english/middle-ground, accessed February 1, 2019.

by motivating our work that is based on empirical findings, and describe the design process of the data transfer system we developed that aims to support existing material exchange practices of ÖGS teachers. On the basis of the given case of the developed data transfer system, we finally discuss how our pursued approach of a *middle-ground* solution (from a researchers' and the technology providers' perspective) may be of value for reflection on existing work practices and a means to articulate issues that are posed by the given context. We elaborate on the potential of indefinite deployment and conclude the discussion by outlining potential further applications of the developed data transfer systems not only within but also outside the educational context.

2 Related Work

As our research touches upon diverse fields of research, this section details related work mainly in two areas: Infrastructuring (e.g., Karasti, 2001; Pipek and Wulf, 2009, Bødker et al., 2017; and Andersson et al., 2018;), and HCI for hearing-impaired children education (e.g., Slegers et al., 2010; Vermeulen et al., 2012, Cano et al., 2016, Korte et al., 2012).

Infrastructuring is a subject that has been given a lot of attention in the CSCW community. The term was first defined by Karasti & Syrjänen (2004) in the context of participatory design and as Bossen et al. (2014) state, infrastructuring focuses on the relation between an infrastructure and its user with issues and dependencies. Although the infrastructure discourse spans multiple domains, our work is situated within the domain work infrastructures (Hanseth & Lundberg, 2001; Pipek & Wulf, 2009; and Stevens et al., 2010). Bødker et al. (2017) argue that most design projects are infrastructuring projects since they build on pre-existing technologies, competencies, and practices; similarly, the work we present in this paper is based on current technologies at the school, the technological competencies of the ÖGS teachers, and their current work practices since.

Following Karasti (2001), it is crucial to initially comprehend work practices and incorporate them in design in order for collaborative technologies to be effective. Infrastructures can be thought as relations that embed choices and politics, meaning that they shall not be perceived as isolated and discrete entities but rather that the use of a single technology emerges in complex relationships and becomes integrated into organisational processes (Andersson et al., 2018). We contribute to the infrastructuring discourse with our work which aligns with existing research, and we are aware of the complex relationships that the development of technology is embedded in and shaped by. However, our contribution diverges in terms of making explicit that this "dedication" to context-sensitivity also leads to a compromise in terms of achieving the *optimal* solution from both a technological and a research perspective; a topic which is rarely reported in scientific work.

As part of our research, the developed system does not only serve the specific purpose of addressing a particular problem; it may be considered a *middle-ground* solution from a research as well as technology providers' perspective, 'embodying' a compromise between the teachers' identified needs for data transfer and the contextual constraints in the school. Consequently, the developed system informs about the complex relationships at the school and serves as a means to *reflect* on given practices that are shaped by given constraints.

Reflection is used as a tool to "think about" since it aids in comprehending and reframing situations leading towards a problem-solving procedure (Schön, 1983). Designing for reflection is not a novel topic; in fact, it has been and is used by researchers in HCI (Baumer, 2015; Baumer et al., 2014; Odom, Banks, Durrant, Kirk, & Pierce, 2012). Drawing on (Fleck & Fitzpatrick, 2010) reflection can serve a multitude of aims (i.e., uncertainty resolution; critical review; reflection on learning); in our case, we use reflection as a tool to create new understanding and appreciation of the said complex relationships formed through the infrastructure of the teachers' work practices but also to communicate those relationships to third parties.

From the perspective of conducting research with and for hearing impaired, Slegers et al. (2010) and Vermeulen et al. (2012) indicate the challenges of pursuing research that involves hearing-impaired children in school contexts, due to their deficiencies in written and spoken language. They suggest following a user-centred design approach to understand the needs of such a vulnerable user group. Kinnula et al. (2018) follow a similar collaborative process in a school setting and introduce a non-context specific, analytical lens for conceptualising, understanding and supporting collaborative design where the value is co-created by the different stakeholders. Morningstar et al. (2015) indicate two dimensions to support inclusion; support for participation and support for learning. In a similar fashion, we contribute to this body of work through our developed system with which we support participation by providing all students access to academic curricula content (sign language vocabulary).

3 The case of addressing a domain-specific problem

This research project's starting point was to collect insights regarding the students' and teachers' communication and learning practices in that particular context, and how technology may facilitate and support them in the future. In our research process (Figure 1) we initially conducted fieldwork in the form of participatory observations and semi-structured interviews and through them, we identified the ÖGS teachers' teaching material exchange practices based on their limited technical solution. The ideal solution, from the researchers' perspective, would involve deploying a state-of-the-art technology readily available on the market or

a technologically advanced tailored platform that supports the teacher's practices (i.e. a digital platform in the form of a smartphone app or a website that hosts a digital sign-language vocabulary with supplementary exercises and a server to save homework). However, given the constraints and complexities of the research context, these options were not feasible since it would require a drastic modification of the teachers and students existing practices. Such modification would subsequently require alteration of the school's technical infrastructure, provision of technical training to support the teachers, and other unpredictable modifications.

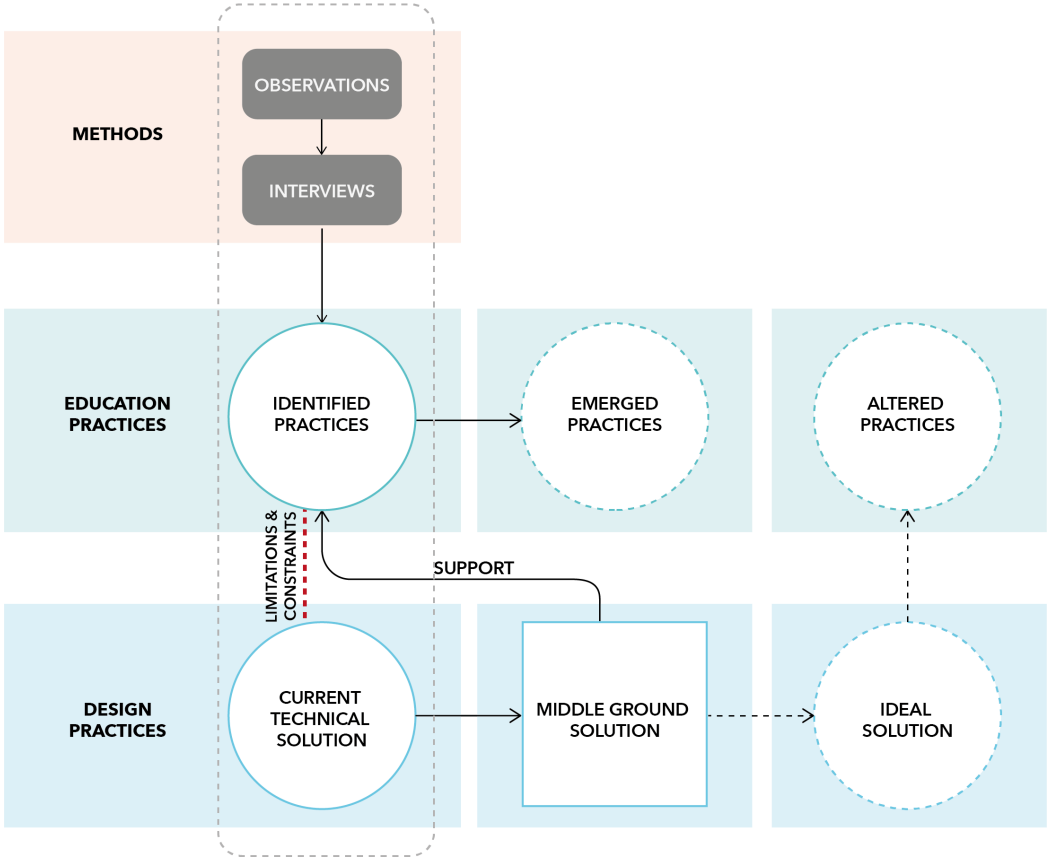


Figure 1. Illustration of the research process we followed based on the utilised methods, the identified teachers' practices and existing technological solutions. The non-dotted shapes indicate steps in the process we already conducted, identified, and developed dotted shapes indicate potential future directions that we could envision to further investigate into.

A less technologically and scientifically advanced solution was selected in favour of supporting existing practices that would simultaneously ameliorate the practices' weaknesses (less time-consuming). We opted for a less drastic modification of the practice which was the most *optimal* option; even though that meant that from a research perspective we had to compromise for a *middle-ground*

solution. In the future, we plan to conduct further fieldwork to discover how the practices have changed based on the introduced solution.

In the following we present the empirical foundations of the work described in this paper, along with an account of the design process of and design rationale for our data transfer system, and the reasons that render the system as a compromise on the researchers' part, in terms of technological possibilities of data transfer procedures.

3.1 Empirical Foundations

In this sub-section we introduce the empirical foundations of our research; the research context and the identified problems. We outline how these insights have informed the design process of the data transfer system we developed. Our system was designed and developed in interwoven phases in constant collaboration with the ÖGS teachers. This section concludes by outlining initial reactions of the teachers to the deployed system.

3.1.1 Research Context

The motivation for this work emerges from the *Diversity-Centred Design* project[5] where we have been, are, and will be collaborating with a number of professionals from various disciplines. The project's aim is to study, analyse, develop, and deliver a better school experience for both the students and the teachers of that specific school.

3.1.2 Data gathering Methods and Data Analysis

In order to learn about the students' and teachers' communication and learning practices, we conducted participatory observations (Flick, 2009). We chose to set-up observational sessions after consulting the school's headmaster who got us in contact with the associated teachers. We (three researchers from the Center for Human-Computer Interaction) observed a Biology lesson, an English language lesson and an Austrian Sign Language (ÖGS) lesson. We made use of observation sheets and schemes (Flick, 2009) to take handwritten notes of technology use, communication and learning practices, and of anything that introduced friction during lesson delivery. The notes were subsequently used as part of the data analysis.

After these observations, we performed three semi-structured interviews (Flick, 2009) consisting of open-ended and more theory-driven questions with the respective teachers of these subjects. Both data sources were textual notes from the observation sheets and notes from the semi-structured interviews that were analysed using an inductive approach to content analysis (Mayring, 2004). We (the

5 https://hci.sbg.ac.at/special-needs/

three researchers that took part of the observations and interviews) familiarised ourselves with the data and individually identified relevant themes. In a joint session, we discussed unclear cases and agreed on the categories. For the purpose of this paper, we only present selected findings that were decisive for the development of our system.

3.1.3 Selected findings

Through our observation sessions, we discovered several issues during class delivery. We clustered those issues in three categories: *spatial, social,* and *technological*. *Spatial* sitting configuration and visibility in classrooms with hearing-impaired children were of high importance, especially in the case of multiple speakers. Similarly, we identified issues regarding *social* encounters in which communication problems seemed to surface quite frequently within the 45-minute lessons; a single class illustrated diverse dynamics in terms of comprehension. The most severe issues seemed to be mainly *technological* ones where teachers would sacrifice time during lessons to attempt to overcome those issues, thus, pausing lesson delivery. Specifically, during the sign language class, we observed that the ÖGS teacher struggled to deliver her lesson, which required that the students performed the new signs they learned that week via recording themselves on Apple iPads placed on their desks; a task that seemed complicated and challenging due to the iPads' low memory capacity.

Based on this observation we wanted to find out more about the teachers and students' practices (and related issues) of video capturing and transfer in sign language classes. During the interview the ÖGS teacher claimed that sign language cannot be taught without the use of technology and indicated that *technological* problems were the biggest issue she was facing during lesson delivery. The teacher gave an account of the way her colleague and herself structured the curriculum of the sign language course on their own, with improvised visual sign language vocabulary booklets (e.g., the teachers taking pictures of one another while signing out and making hand-written notes below to translate the sign) and self-recorded videos (e.g., created at home with an iPad on a tripod) as there are non-existent guidelines to follow nor ready-made material they could use. This is due to a very small Austrian sign language community and an even smaller number of teaching professionals. The ÖGS teacher explained how *data transfers* and specifically, *distributing* video recordings through USB flash drives to the students is crucial, not only to teach them new signs but also to assign homework. In return, the second data exchange takes place when the children have to record themselves at home performing the words the teacher assigned, and save the recording onto a removable data storage device such as USB flash drives. One of the issues the teacher mentioned was that a number of children do not have access to a personal computer with a web camera at home or even an internet connection. During class, the teacher

copies all the homework from each child's USB so that she can watch all the videos and grade them in her free time.

3.2 The case of an off-web data transfer system for teaching material exchange practices

Based on the issues we identified throughout our empirical fieldwork, we saw the opportunity to develop a technical solution that would support the ÖGS teachers in their everyday practice of exchanging teaching materials with the students, since we observed that they were relying the most on technology during lesson delivery. The solution is part of a bigger project that revolves around a series of design interventions we will deploy at the school, that aim to address identified teachers' needs. Additionally, any teacher working with assorted media at the school can potentially benefit from our developed solution, e.g. for exchanging files with other teachers, or other media files with students. In this section we will describe the design process we have followed; the initial concept, the early prototyping that led to low-fidelity prototypes, and we will conclude by describing the developed system that aims to support the teachers' practices in terms of functionality and components.

3.2.1 Initial concept and scenario development

A first concept diagram was drafted to communicate to other members of the research team what the teacher and the interviewer had discussed as an initial solution as part of the semi-structured interviews. Specifically, the teacher requested a system that she could plug-in all USB drives and it would perform the copying process to and from all the students' USB drives as simple as possible. The diagram, as illustrated in Figure 2, comprises of the *two scenarios* of data transfer use via a stand-alone device; (1) data transferred from the teacher's USB drive to the children, and (2) vice versa. A button switch is assigned to each scenario of use.

- Scenario 1: The students copy their homework files onto their USB drives in the 'homework' folder. At the beginning of the lesson, the students connect their USB drives to their designated USB port on the system. A green button press transfers the homework files from the children's USB drives to the teacher's.

- Scenario 2: This scenario involves the teacher updating the vocabulary files onto her USB drive. Then, with the press of a blue button switch, the files are transferred to all children's USB drives, in a 'vocabulary' folder.

The initial idea was extending the functionality of an existing USB hub by connecting it on a micro-controller. An additional USB port connected directly to

the micro-controller would act as the teacher's USB port. All the components would be enclosed in a fabricated acrylic casing.

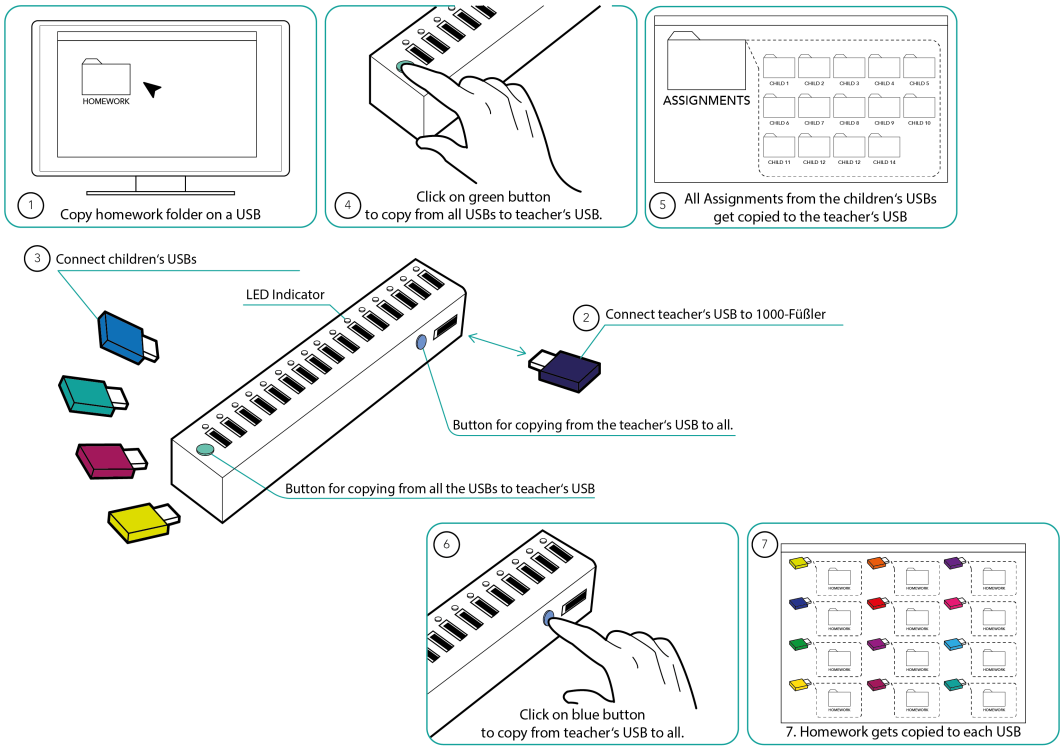


Figure 2. Concept drawing of the data transfer system.

3.2.2 Early Prototyping

Following specifics from the teacher, such as the wish to have a device smaller than a standard A4 page size and slimmer than a book so that it can fit in her backpack, we commenced exploring initial physical forms for the system using cardboard modelling (Figure 3). In designing the cardboard mock-ups, the team took a set of design decisions; ensuring the device's size would be able to fit in a backpack and that no sharp edges should be exposed. The next step was creating a proof of concept by coding the behaviour of the single-board computer and testing our ideas in practice.

3.2.3 Low-fidelity prototype and initial feedback cycles

A low-fidelity prototype was developed via connecting a self-powered USB hub and two button switches to a single-board computer (Raspberry Pi 3 model B). We managed to transfer data from one primary USB drive to four other that were connected on the USB hub and vice-versa through triggering the push switches. In order to communicate the status of the device to the user, we assigned four LED colours; i.e., green indicated that the system is ready for the data transfer, blinking

blue indicated data transfer in progress, constant blue indicated successful data transfer, yellow indicated a data transfer error, and red indicated a system error. A casing fabricated out of white acrylic (selected due to durability) enclosed the hardware components.

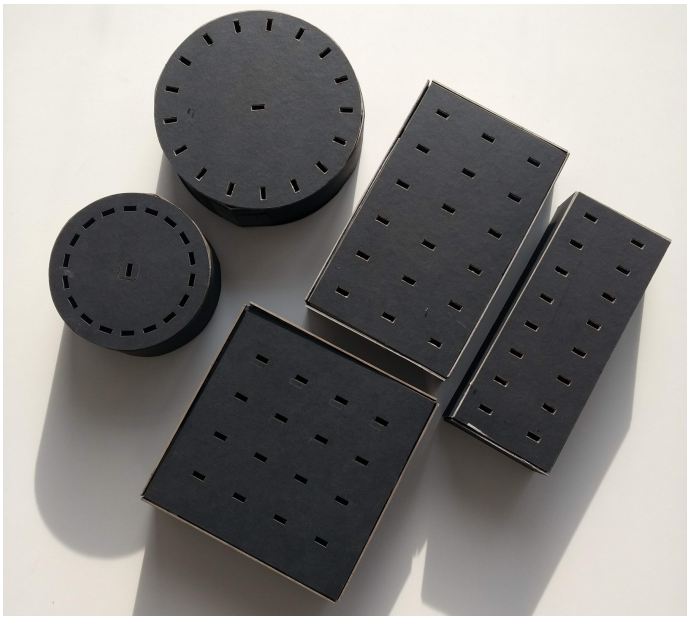


Figure 3. Photograph showing the five initial form explorations in black cardboard.

We invited the ÖGS teacher for a meeting, where we presented the tangible representation of the concept we had discussed during the initial interview. During our demonstration of the early prototype (Figure 3), the teacher embraced our ideas and suggested minor changes (e.g., the USB slots configuration, the form factor, and the folders' organisation format). We discussed the system's specifics regarding functionality related to her practice (i.e., the folder organisation system she's using) to align the two as much as possible.

3.2.4 Final design of the data transfer system

The end result of our design process was a stand-alone data transfer system (Figure 4) that consisted of a square casing with 19 USB hubs (one for the teacher and 18 for the students) with corresponding LEDs and two buttons to copy one or several data files from either a teacher's USB drive (source) to multiple USB drives (target) and vice versa[6]. In order to reduce the number of interactions to the bare minimum, whilst making sure that the system is comprehensible, we automatised repetitive tasks and implemented a light pattern for signalling the device's status. Our system now allows for a once a week homework assignment instead of a biweekly one,

[6] Video-demonstrator of the final data transfer system and its functionality. https://myfiles.sbg.ac.at/index.php/s/XOEKIPnhKizmEgo/download

since the teacher does not have to take the students' USB drives home to perform the data exchange.

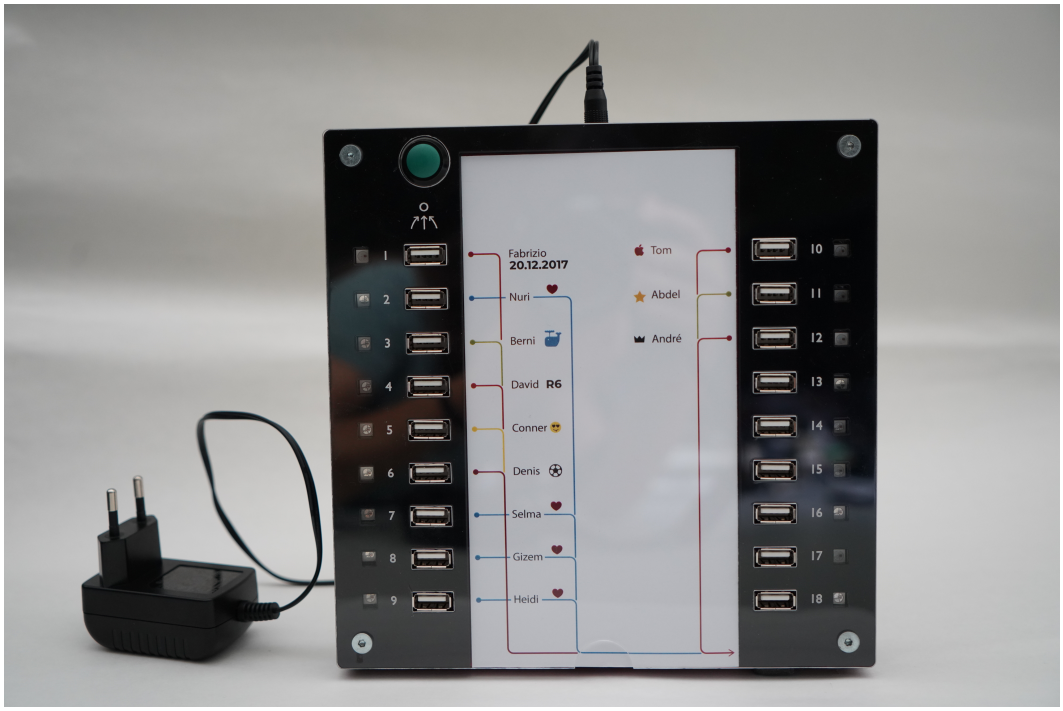


Figure 4. The deployed data transfer system features a slot that holds interchangeable keys (tailored to each class' name list) that assign each USB port to a student.

Hardware Components

During lesson delivery, all homework files are copied to her USB drive, the updated vocabulary files and next week's homework are copied to all students' USB drives.

Consequently, two 10-port USB 2.0 Hubs were employed in combination with a Raspberry Pi Zero W (to minimise the physical size) for building the main functionalities of our system. Moreover, we added a LED strip to indicate the aforementioned different states of each port. All components of the system are powered through a central 5V / 2A power supply. The software was written in Python and made use of different libraries to be able to interact with the connected USB hubs and LED-Strip.

The thickness of all components was calculated in order to build a casing as small and slim as possible. We laser-cut acrylic layers with cut-outs to the exact size of the internal components, stacked them and mounted them with screws and bolts to create a solid casing.

Functionality of the data transfer system

When an USB drive is inserted into a port and recognised by the system, the corresponding LED turns green. The minimum copying requirements are that the

teacher's USB drive and one of the students' USB drives are plugged-in and that the student drives contain the designated folder for homework (named "ABGABE"). If these prerequisites are fulfilled, copying is possible with the press of either, a) the blue button, for copying from teacher-to-students, or b) the green button, for copying from students-to-teacher. Following a button press, the corresponding LEDs of the plugged in USB drives commence to blink successively and then light in constant green. This pattern indicates that the system is mounting all drives.

Once all LEDs light up in constant green, the system initiates the copying procedure. As soon as copying is finished, the LEDs turn blue. As soon as the LEDs turn constant blue, the corresponding USB drive can be safely removed Respectively, in case of a copying error, the LEDs turn yellow, while in case of a system error, one or more LEDs turn red.

4 Discussion and Outlook

In the following, we discuss the outcome of our approach in terms of applying the *middle-ground* solution and its potential. We further elaborate on the case of the data transfer system as a means for reflection and articulation of issues, we refer to the implications of its indefinite deployment, and we elaborate on its further applications.

4.1 The *middle-ground* solution

One of the most striking discussion points that emerged through our research and design approach was the twofold way of reading the solution. Third parties (e.g., other project partners involved in our research project) immediately identified other state-of-the-art technology that would support the teachers' material exchange practices (i.e., USB data transfers) and, therefore, considered them as a workaround to deal with the constraints of given contexts (e.g., limited technological infrastructure in the school and the students' homes). In contrast, the ÖGS teachers considered their practices as the only way to operate, and *they* perceived the developed system as the *ideal* solution. We argue that even though there are widespread technological and scientific solutions that could pose as the better option (i.e. a technologically advanced platform that could serve as an online dictionary or a smartphone vocabulary application, while additionally, a local server could host the students' assignments), the case of the data transfer system serves, from a research perspective, as the *middle-ground* solution that supports the teachers' practices and alleviates issues they were facing during lesson delivery. If a technology-wise *ideal* solution in the form of a non-web-based data transfer platform or a cloud-based data transfer system was employed, it would unavoidably

require alteration of the teacher's practices and the school's current technological infrastructure as well as teacher's training of how to use the new system.

Aiming for the *middle-ground* solution is an unusual but not a novel approach; it has been utilised as a management strategy for software development as the most optimal approach based on total cost minimisation (Goldstein et al., 2010). Indeed, as we experienced in the case of the data transfer system, employing the *middle-ground* solution in certain contexts might be more optimal than chasing the *ideal* solution. This resonates with work by Pipek and Wulf (2009), as e-infrastructures also follow the same logic; their employment could improve the design of IT infrastructures in organisations.

We strongly believe that the reported approach has the potential of becoming a technique. Through the case of the developed data transfer system, we explored how a *middle ground* solution could be the optimal one for the teachers based on the limitations posed by the given context. However, conforming to Wulf et al. (2011), design case studies in other fields of practice are necessary in order to identify cross-cutting issues to compare and combine insights from those cases. The pursued approach in this research represents a highly contextual and user-centred one. For future work we aim to further elaborate on how such middle ground solutions may serve as highly contextualised indicators of (technological, infrastructural, or else) change; meaning that indefinitely deploying such solutions and accompanying this deployment over time may help to identify new opportunities for technological innovation.

4.2 Reflection and articulation of issues

Apart from the practical contribution of a fully functioning data transfer system, this system may be furthermore considered as a means to *reflect* on and exhibit the problems such a school might face. We envision to use this system as a reflection, to initiate discussion and demonstrate the issues that the ÖGS teachers were facing, the solution we developed and what would have been the *ideal* solution to third parties (i.e., future stakeholders). Specifically, through the mere existence of the developed system we reflect on the unreliable technical infrastructure that does not allow for smooth lesson delivery. Through this reflection process we are attempting to reframe the situation to relevant parties which could possibly lead towards a problem-solving procedure on a higher level (Schön, 1983).

In addition to a reflection, the system serves as a supplementary articulation means that illuminates prevailing yet overlooked issues to present them to teachers, parents, and other potential future stakeholders (e.g., representatives from the deaf community). It may be read as a kind of tangible representation of the practices and workarounds the teachers have to perform on a daily basis 'to get their work done'; it embodies deficiencies in the educational system (e.g., lacking professional teaching curriculum for Austrian sign language), deficiencies in sharing and

exchanging knowledge as part of a community of practice, or deficiencies in the current infrastructure of the school (e.g., lack of internet access) or the students home (e.g., partially not having full internet access or computer access).

4.3 Indefinite deployment of a *research product*

This subsection refers to the indefinite deployment of the developed research product, independent from conceptualising this research product as an ideal or middle-ground solution. As argued by Lim et al. (2008), the fields of HCI, software engineering, and design commonly use the term 'prototype' to indicate that a certain artifact is used as part of a design process. In these fields, the importance of prototypes is obvious and unquestionable. However, as argued by Odom and colleagues (2016), diverse complexities and challenges emerge when researching human-technology relations in *real-life* contexts over a *longer period of time*. 'Prototypes' may not be sufficient enough to research questions related to these complexities (Odom et al., 2016). *"While the fidelity of prototypes can range, they remain references to future products, systems, or services*" (Odom et al. 2016, p. 2549). In this perspective, prototypes may be considered as placeholders for something else; an instantiation of a *future outcome* (Lim et al., 2008). In this line of thought, Odom and colleagues (2016) suggest that the concept of a 'prototype' might not be adequate to support inquiries regarding everyday life and introduce the notion of *research products* whose explicit aim is actuality (i.e., users experience the artifacts as they are and not what they might become*)*. With our developed system, we have created a research product that can be indefinitely deployed in real-life context, without any dedicated maintenance from our side; meaning it can be used *as is*, rather than what it *might become*.

Our work does not only contribute a fully functioning system to support teachers' work practices but also, a research product that is deployed in the school in an indefinite manner and, is therefore, also open for future explorations and research. This is of particular importance, as prototypes developed in HCI and CSCW research, are often no longer in working order or even existence. Work that reports on designs or relevant deployment studies which make use of research prototypes is often archived, but the said prototypes are not (e.g., Truong et al., 2015).

Truong et al. (2015) point out several factors that define the length of a system deployment in the wild such as the context, frequency of use, and shelf-life. In contradiction, our deployed system stands out by providing the teachers with a solution that they will use in their everyday practices indefinitely (as long as the users wish to) without the researchers collecting use data, iterating on the system, or retiring it back to the research facilities. Through the established trust and collaboration between researchers and ÖGS teachers, a direct communication channel is in place in order to revisit the school to see how our system facilitates and/or alters the teachers' initial work practice. The importance of long-term

deployments has been argued for a long time, due to several benefits. According to Karapanos et al. (2009), prolonged use of a deployed system allows for meaningful mediation. Moreover, a confluence between the said system and the pre-existing work practices of that context is crucial in order for the introduced system to be utilised in a meaningful way (Pipek and Wulf, 2009). Additionally, during long deployments, interesting relationships might unfold among people and computational things (Odom et al., 2016).

Deriving from our gathered insights, there is a set of design attributes to take into consideration when designing technology for indefinite deployment such as: the life-span of the electronic components based on the frequency and purpose of use, the high quality of finish, the material durability of both the external casing and the hardware components and a plan for unexpected maintenance, such as, easy access to electronic components. However, caution should be given when deploying a solution indefinitely due to the disadvantages it might hinder, such as slowing down solution development or solution stagnation. In our view, researchers should monitor indefine deployment; in case of emergent opportunities for further technological development (e.g., changes to the technological infrastructure) that would allow for further iterations on the initial solution or the development of an altogether new solution.

4.4 Further Applications for the data transfer system

As far as further applications go, other teachers working with assorted media can benefit from the developed system (e.g., digitally distributing or collecting homework with and from students). Additionally, we are confident that open-sourcing our system on a platform that hosts do-it-yourself projects such as Github[7] or Instructables[8] will increase its availability. Thus, other individuals facing similar technological limitations and simultaneous data transfer needs can fabricate their own system without purchasing any sophisticated equipment. Purchasing the bare minimum material requirements of two ordinary USB hubs, a raspberry pi single-board computer, a common LED strip, and some basic coding skills would suffice, rendering the system into widely available equipment. Following the use of the system on a daily basis since deployment (five months), the ÖGS teacher to whom the system was delivered, benevolently suggested to make the system available to other sign language teachers in other country regions; feedback that reassured us of the system's appropriateness and its widespread applications.

We envision that a data transfer system could prove useful in contexts where people meet, such as events where quick data transfers are called for (i.e.,

7 Github is web-based hosting platform for software development projects. https://github.com

8 Instructables is web-based platform that hosts user-created and uploaded do-it-yourself projects http://instructables.com/

international conferences or consortium meetings where exchanging general info simultaneously is necessary). Furthermore, in offices where sharing files wirelessly or using web-based services is not an option or in settings where confidentiality is an issue and keeping files off-line is of great importance, the system could be used as the main data transfer method. Moreover, we see this system utilised for recreational and entertainment purposes (i.e., sharing music or video content, art installations, design exhibitions).

5 Conclusion

Domain-specific problems encountered at work practices may seem insurmountable to the end user. In this paper, we have reported on our research project at an integration school with hearing and hearing-impaired staff and students where we observed the work practices of the sign-language teachers. Through our empirical insights, we have identified domain-specific problems within the said practices; explicitly, disruption of lesson delivery due to mundane, time-consuming teaching material exchange procedures. This paper contributes to the CSCW community through presenting an approach for addressing domain-specific problems with a *middle-ground* solution. In addition, we have reported on and contributed the resulting outcome of this approach, the case of a data transfer system that supports the sign-language teachers' practices and alleviates their issues. The developed system can be read as a tangible representation of the contextual issues the sign-language teachers are facing and as a means of a reflection on current practices and an articulation of the limitations that constraint them. We have reflected on our approach and discussed the resulted case by outlining potential new roles for research products, elaborating on long-term deployment, and speculating on alternative application domains. As demonstrated, *middle-ground* solutions can be the most optimal ones in supporting work practices instead of chasing after the ideal. As the *ideal* solution, from a research and technology advancements' perspective, might not always be feasible or accessible, we argue, that there is a need to also advance middle-ground systems. Researching what might not be cutting-edge (from a technological perspective) or ideal (from a users' perspective) requires specific research and design approaches we aim to further advance in future research.

6 Acknowledgments

Our sincere thanks to the students, the teaching staff, as well as the headmaster of the Josef Rehrl School who gave us permission to observe their practices and discussed with us about their work that led to a naturally evolved collaboration. This research is part of a joint research endeavour in partnership with the Salzburg Wohnbau GmbH, the Salzburg AG, and the Josef Rehrl School.

7 References

Andersson, C., Cozza, M., Crevani, L., & Schunnesson, J. (2018). Infrastructuring for remote night monitoring: frictions in the strive for transparency when digitalising care service. *Proceedings of the 16th European Conference on Computer-Supported Cooperative Work - Exploratory Papers*, *2*. https://doi.org/10.18420/ecscw2018_6

Baumer, E. P. S. (2015). Reflective Informatics: Conceptual Dimensions for Designing Technologies of Reflection. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, 585–594. https://doi.org/10.1145/2702123.2702234

Baumer, E. P. S., Khovanskaya, V., Matthews, M., Reynolds, L., Schwanda Sosik, V., & Gay, G. (2014). Reviewing reflection: on the use of reflection in interactive system design. *Proceedings of the 2014 Conference on Designing Interactive Systems - DIS '14*, 93–102. https://doi.org/10.1145/2598510.2598598

Bødker, S., Dindler, C., & Iversen, O. S. (2017). Tying Knots: Participatory Infrastructuring at Work. *Computer Supported Cooperative Work (CSCW)*, *26*(1–2), 245–273. https://doi.org/10.1007/s10606-017-9268-y

Bossen, C., Ehn, P., Karasti, H., Di Salvo, C., Clement, A., Pipek, V., & Dittrich, Y. (2014). Infrastructuring, collaboration and evolving socio-material practices of changing our world. *Proceedings of the 13th Participatory Design Conference on Short Papers, Industry Cases, Workshop Descriptions, Doctoral Consortium Papers, and Keynote Abstracts - PDC '14 - Volume 2*, 221–222. https://doi.org/10.1145/2662155.2662211

Cano, S., Collazos, C., Fardoun, H. M., Alghazzawi, D. M., & Albarakati, A. (2016). Model Based on Learning Needs of Children with Auditory Impairment. In G. Meiselwitz (Ed.), *Social Computing and Social Media* (pp. 324–334). Springer International Publishing.

Fleck, R., & Fitzpatrick, G. (2010). Reflecting on reflection: framing a design landscape. *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction - OZCHI '10*, 216. https://doi.org/10.1145/1952222.1952269

Flick, U. (2009). *An introduction to qualitative research* (4th ed). Los Angeles: Sage Publications.

Goldstein, Z., Petrie, D., & Sherif, Y. (2010). Finding middle ground management strategy for software development. *Information Management & Computer Security*, *18*(3), 185–197. https://doi.org/10.1108/09685221011064717

Hanseth, O., & Lundberg, N. (2001). Designing Work Oriented Infrastructures. *Computer Supported Cooperative Work (CSCW)*, *10*(3–4), 347–372. https://doi.org/10.1023/A:1012727708439

Ideal. (n.d.). In *English Oxford Living Dictionaries*. Retrieved from https://en.oxforddictionaries.com/definition/ideal

Karapanos, E., Zimmerman, J., Forlizzi, J., & Martens, J.-B. (2009). User experience over time: an initial framework. *Proceedings of the 27th International Conference on Human Factors in Computing Systems - CHI 09*, 729. https://doi.org/10.1145/1518701.1518814

Karasti, H. (2001). Bridging Work Practice and System Design: Integrating Systemic Analysis, Appreciative Intervention and Practitioner Participation. *Computer Supported Cooperative Work (CSCW)*, *10*(2), 211–246. https://doi.org/10.1023/A:1011239126617

Karasti, H., & Syrjänen, A.-L. (2004). Artful infrastructuring in two cases of community PD. *Proceedings of the Eighth Conference on Participatory Design Artful Integration: Interweaving Media, Materials and Practices - PDC 04*, *1*, 20. https://doi.org/10.1145/1011870.1011874

Kinnula, M., Iivari, N., Isomursu, M., & Laari-Salmela, S. (2018). 'Worksome but Rewarding' – Stakeholder Perceptions on Value in Collaborative Design Work. *Computer Supported Cooperative Work (CSCW)*, *27*(3–6), 463–494. https://doi.org/10.1007/s10606-018-9328-y

Korte, J., Potter, L. E., & Nielsen, S. (2012). Designing a Mobile Video Game to Help Young Deaf Children Learn Auslan. *BCS-HCI '12 Proceedings of the 26th Annual BCS Interaction*

Specialist Group Conference on People and Computers, 345–350. Swinton, UK: British Computer Society.

Lim, Y.-K., Stolterman, E., & Tenenberg, J. (2008). The anatomy of prototypes: Prototypes as filters, prototypes as manifestations of design ideas. *ACM Transactions on Computer-Human Interaction*, *15*(2), 1–27. https://doi.org/10.1145/1375761.1375762

Mayring, P. (2004). Qualitative Content Analysis. In *A Companion to Qualitative Research* (Uwe Flick, E.v. Kardorff, and I. Steinke, p. 144). London: SAGE.

Middle ground. (n.d.). In *English Oxford Living Dictionaries*. Retrieved from https://en.oxforddictionaries.com/definition/middle_ground

Morningstar, M. E., Shogren, K. A., Lee, H., & Born, K. (2015). Preliminary Lessons About Supporting Participation and Learning in Inclusive Classrooms. *Research and Practice for Persons with Severe Disabilities*, *40*(3), 192–210. https://doi.org/10.1177/1540796915594158

Odom, W., Banks, R., Durrant, A., Kirk, D., & Pierce, J. (2012). Slow technology: critical reflection and future directions. *Proceedings of the Designing Interactive Systems Conference on - DIS '12*, 816. https://doi.org/10.1145/2317956.2318088

Odom, W., Wakkary, R., Lim, Y., Desjardins, A., Hengeveld, B., & Banks, R. (2016). From Research Prototype to Research Product. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, 2549–2561. https://doi.org/10.1145/2858036.2858447

Pipek, V., & Wulf, V. (2009). Infrastructuring: Towards an Integrated Perspective on the Design and Use of Information Technology. *Journal of the Association of Information Systems*, *10*(5), 306–332.

Schön, D. A. (1983). *The Reflective Practitioner: How Professionals Think In Action*. London: Temple Smith.

Slegers, K., Duysburgh, P., & Jacobs, A. (2010). Research methods for involving hearing impaired children in IT innovation. *Proceedings of the 6th Nordic Conference on Human-Computer Interaction Extending Boundaries - NordiCHI '10*, 781. https://doi.org/10.1145/1868914.1869028

Stevens, G., Pipek, V., & Wulf, V. (2010). Appropriation Infrastructure: Mediating appropriation and production work. *Spec. Issue EndUser Development. Int. Journal of Organizational and End User Computing (JOEUC)*, *22*(1).

Truong, K., Kientz, J., Banerjee, N., Brush, A. J., & Mahajan, R. (2015). Deployment Study Length: How Long Should a System Be Evaluated in the Wild? *GetMobile: Mobile Computing and Communications*, *19*(2), 18–21. https://doi.org/10.1145/2817761.2817768

Vermeulen, A., De Raeve, L., Langereis, M., & Snik, A. (2012). Changing Realities in the Classroom for Hearing-Impaired Children with Cochlear Implant. *Deafness & Education International*, *14*(1), 36–47. https://doi.org/10.1179/1557069X12Y.0000000004

Wilcox, S., & Occhino, C. (2016). *Historical Change in Signed Languages*. https://doi.org/10.1093/oxfordhb/9780199935345.013.24

Wulf, V., Rohde, M., Pipek, V., & Stevens, G. (2011). Engaging with practices: design case studies as a research framework in CSCW. *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work - CSCW '11*, 505. https://doi.org/10.1145/1958824.1958902

Schwaninger, I.; Fitzpatrick, G.; Weiss, A. (2019): Exploring Trust in Human-Agent Collaboration. In: Proceedings of the 17th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing an the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep08

Exploring Trust in Human-Agent Collaboration

Isabel Schwaninger, Geraldine Fitzpatrick, Astrid Weiss
Vienna University of Technology
{isabel.schwaninger, geraldine.fitzpatrick, astrid.weiss}@tuwien.ac.at

Abstract. Collaborative ensembles will increasingly involve agents like robots in the near future. A key part of collaboration is trust. While trust has been mainly studied between humans in CSCW, trust in human-agent research has been mostly studied in dyadic formations divorced from the broader context. This exploratory paper critically discusses previous work on trust across CSCW and HCI-related areas, taking into account recent practice approaches in CSCW and what they can contribute for understanding trust in human-agent collaboration. To make better sense of how trust emerges in collaborative ensembles with agents, we suggest that concepts that have been proposed in the field of human-agent interaction need to be further explored in real-life settings, while concepts embraced in CSCW can lead to a more thorough understanding of the situatedness and dynamics of trust going beyond the attributes of the agent itself.

1 Introduction

Understanding key aspects of collaboration is an important part of CSCW and HCI research. In the near future, due to technological innovation, we can anticipate that collaborative ensembles will increasingly involve agents like robots. For example, as assistive robots are developed for care, collaborative care networks may involve at least patients, caregivers and robots. An agent, like the robot here, can be defined

as "an object or technology that people interact with as if it is able to act with its own purposes, motivations, and intentions"[1] (Human-Agent Interaction, 2019).

Trust is a key aspect of collaboration, and it has been studied in both CSCW and in other research with agents. In the CSCW community, trust has been studied mainly with regards to human-human collaboration. In this context, technology is a tool through which interpersonal trust can be fostered, without being an artifact or even a perceived social actor to be trustworthy itself. In other communities, such as in Human-Agent Interaction (HAI) and Human-Robot Interaction (HRI), the agency and social role of agents has been taken into account. Given that trust is a key aspect of collaboration, the aim of this paper is to understand how trust is differently framed across disciplines, and what CSCW can contribute to trust in human-agent collaboration.

We consider human-agent collaboration, in accordance with Jung et al. (2018), as collaborative partnerships between humans and agents in completing tasks that typically focus on coordinating close, seamless joint activities. However, as also mentioned by Jung et al. (2018), up to now this research often only focused on human-agent dyad studies, but interaction scenarios are becoming more and more common in which one or several robots are deployed in social contexts that involve groups of people rather than individuals. In the simplest of collaboration scenarios that involve multiple people, an agent will have to make decisions about how to distribute resources (e.g. social attention, task support, or physical resources) and preliminary studies in HRI have already studied the effects of different types of collaboration formations on human-robot teaming (Brosnan and de Waal, 2014; Lee, 2018). Different collaboration formations are also likely to impact issues of trust in human-agent collaboration studies but to date this has not been well explored, especially in complex real world settings and collaborative formations.

In this exploratory paper, trust is critically reflected on as it is used both in CSCW and in other work on human-agent collaboration. Conceptualizations of trust as they are used in human-agent research are critically discussed, taking into account recently emerging practice approaches in HCI. Our aim is to outline broad directions for future research, which can then form the basis from which future studies can be defined and conducted.

The outline of this paper is as follows. In the next section, previous work on trust in CSCW and HCI will be critically reflected on. Subsequently, a section elaborating work on trust in agents will discuss agents as social actors and previous work on trust in human-agent interaction. We then discuss what building blocks are missing that need further research, and the contributions that CSCW can make to understanding and developing better human-agent collaboration, followed by a Conclusion section.

[1] http://hai-conference.net/what-is-hai/

2 Trust in CSCW

In CSCW, and related Human Computer Interaction (HCI) research, trust has been studied in various contexts and mainly on an interpersonal level. Technical artifacts can be seen as tools entailed in collaborative work or as mediums through which that work is conducted. In this context, interpersonal trust is either framed as a key part of collaboration between people, or as occurring between individuals and political or corporate institutions. For the most part, collaboration is interwoven with its broader social context.

Trust has been particularly identified as important in computer-mediated communication (CMC) (Zheng et al., 2002) and virtual work teams (Al-Ani et al., 2013b; Robert, 2016; Quan-Haase and Wellman, 2005; Bos et al., 2001). In a study on globally distributed computer-supported work, trust was framed as expectations of other human parties (Al-Ani et al., 2013a). Individuals' baseline trust and its effect on the diffusion of trust in cooperation has been explored (Wang and Redmiles, 2016), as well as inter-group trust formation (Nguyen and Canny, 2007) and interpersonal conflict in technologically-mediated settings (Billings and Watts, 2007). In computer-supported crisis management, psychological and social factors were taken into account, and trust was related to information sharing behaviour in a crisis response system (Linot, 2018). Lampinen et al. (2016) have also worked on trust in the context of the sharing economy.

Besides collaboration in work teams, institutional trust has been studied, such as in e-governance and related to e-participation (Corbett and Le Dantec, 2018; De Cindio et al., 2007). Further, Wang and Mark (2013) explored trust in online news, where they compared social media to official news to study trust in institutional practices. Other interesting work showed how people in a political conflict zone were able to create a context of trust (Semaan and Mark, 2011). The use of ICTs helped people to manage their public identity, to conduct background checks, and to develop collaborative practices. Social interaction through technology added to the formation of a context of trust, where trust was framed as a practice. Trust in an institution was also studied in the context of e-commerce (Kim et al., 2017; Greenspan et al., 2000; Egger, 2000; Garnik, 2004), where trusting an online supplier is a crucial part of the trust relationship. In related work, cultural backgrounds have been taken into account (Garnik, 2004), and prepurchase knowledge besides interface properties and informational content (Egger, 2000), as well as interpersonal cues to measure affective reactions related to trust (Riegelsberger, 2003). Other related work has linked privacy and trust (Crabtree et al., 2017; Hong, 2009). E.g. Crabtree et al. (2017) conducted an ethnomethodological study of digital privacy practices in homes, and they found that people were concerned with the impact of the networked world on interpersonal affairs in they daily lives.

Besides a focus on trust in humans or institutions, properties of computers have been studied and how they can also foster interpersonal trust. For example, trust can be affected by choices in the design of a web interface (Marsh and Meech,

2000), and Kostakos and Oakley (2009) explored this through using locative images. Also, design principles have been proposed to foster trust between interacting human parties in collaborative work, drawing from ethnographic fieldwork (Knowles et al., 2015).

In CSCW, contextual knowledge and practices have been taken into account as essential for trust, and this has methodological implications. The use of qualitative research methods such as ethnography and ethnomethodology is thus common in the study of trust, also with an increasing focus on 'practices'. While early methods in HCI were inspired by psychological sciences involving controlled short-term, lab-oriented studies, which are according to Kuutti and Bannon (2014) embedded in the *Interaction paradigm*, this is not the case in the recently emerging *Practice paradigm*. In previous practice-oriented work, the practical accomplishment and "dynamic and situated 'interactional' aspects [...] to be accounted"(Fitzpatrick, 2003, p. 91) was highlighted. Generally speaking, practice approaches explore "[...] historical process and performances, longer-term actions which persist over time, and which must be studied along the full length of their temporal trajectory[,][...] situated in time and space"(Kuutti and Bannon, 2014, p. 3543). Further, the broader context is taken into account, and it is "intervowen within the practice" (Kuutti and Bannon, 2014, p. 3543).

Qualitative studies have shown how trust is enacted through ongoing practices, where it is for example operationalized by public officials (Corbett and Le Dantec, 2018). Trust as a practice itself has been worked on in CSCW already in the early days (Van House et al., 1998), being one of many representatives of the *turn to practice* (Kuutti and Bannon, 2014) in HCI.

3 Trust in Agents

Trust has also been explored in situations involving agents. As agents are becoming part of of collaborative ensembles, previous research has shown that some agents can be treated and seen as social beings (Coeckelbergh, 2012). In the process of anthropomorphizing agents, some authors argue that it is possible to associate human-like characteristics to agents such as benevolence, competence and integrity. From this point of view, an agent can be perceived as a social actor (Waytz et al., 2014).

Based on the assumption that agents can be perceived as social, studies have explored how interpersonal trust occurs in human-robot interaction. In short-term child-robot interactions for example, van Straten et al. (2018) showed that interpersonal trust occurred with robots, where interpersonal trust was distinguished by the children from technological trust. Technological trust can be defined as "the attitude that an agent will help [to] achieve an individual's goal in a situation characterized by uncertainty and vulnerability" (Lee and See, 2004, p. 51). Further, an agent's technical properties was one of the reasons for children to trust robots interpersonally (van Straten et al., 2018).

When reviewing the literature of human-agent collaboration, there are two integrative models of trust by Mayer et al. (1995) and Rousseau et al. (1998) that mostly occur (e.g. Martelaro et al., 2016). In the Mayer et al. (1995) model of trust, trust is defined as the "willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party" (Mayer et al., 1995, p. 712). Trust arises referring to another party's ability (i.e. competence), integrity (i.e. adherence to honesty and truthfulness), and benevolence (i.e. the willingness to protect and support without an egocentric motive) (Mayer et al., 1995). The model of trust by Rousseau et al. (1998) gives a definition for organizational contexts that comes up slightly later, where the notion of social trust is defined as a "psychological state comprising the intention to accept vulnerabilities based upon positive expectations of the intentions or behaviour of another" (Rousseau et al., 1998).

Human-agent trust was modeled by Gulati et al. (2018), who take several properties of both the agent the person into account. A person's motivation and willingness are identified as factors having an impact on trust, along with the perceived agent's competence, benevolence, honesty, predictability and reciprocity. In a study with Siri, they show that trust can be affected by how helpful (or benevolent) an agent is, how competent (or reliable) it is, and how reciprocal it is to a person's needs. However, trust is not significantly affected by motivation, predictability and honesty (Gulati et al., 2017).

Although focusing on dyadic relationships between trustor and trustee is very common, broader approaches have been proposed as well, i.e. on trust as reliance. Billings et al. (2012) proposed a three-factor model of trust in robots, including *human* characteristics such as ability and personality, *environmental* characteristics such as task and team, and *robot* characteristics such as performance and attributes (Billings et al., 2012). These three factors have also been identified in a meta-analysis on trust (Hancock et al., 2011), where the authors stressed that too few studies have yet been conducted on environmental and human-related factors, although robot-related factors have been shown to affect trust the most. The conclusion that robot-related factors are more "influential" on trust is however not convincing, given the few studies on human-related and environment-related factors, where the impact is not yet known. Moreover, identifying and adding the interplay of these factors on one another for trust is still open to research.

Drawing on the model from organizational contexts by Mayer et al. (1995) and the model on trust in automation by Lee and See (2004), Wagner et al. (2018) propose a trust model based on risk. They define trust as "a belief, held by the trustor, that the trustee will act in a manner that mitigates the trustor's risk in a situation in which the trustor has put its outcomes at risk" (Wagner et al., 2018, p.26:4). Trust is modeled in game-theoretic terms, and similar to what Hancock et al. (2011) proposed, they highlight three important factors that influence trust-based decisions, namely the trustee, the trustor, and the situation. The model was also tested in an emergency experiment by Robinette et al. (2016), where

people tended to overtrust the robot despite half of them observing the same robot performing poorly in a navigation guidance task minutes before.

Based on the three-factor model by Hancock et al. (2011), Hoff and Bashir (2015) have also suggested a three-layered model in which these factors contribute to *dispositional*, *situational* and *learned* trust (Hoff and Bashir, 2015). They pointed out that age, gender, culture and personality differences are components of dispositional trust, where they reflect an individual's overall tendency to trust in automation. Situational trust is shaped by internal and external variability, such as self-confidence, mood, type of system, perceived risks, task difficulty and organizational setting. Learned trust includes an operator's evaluation of a system based on past experience or the current interaction, and it is shaped by preexisting knowledge, system performance and design features. The three factors add to a person's reliance on a system, where they can change with the course of a single interaction (Hoff and Bashir, 2015). According to this model, trust affects a person's willingness to rely on a system, which can potentially change through dynamically learned trust, i.e. interaction experience. The latter is affected by system performance and design features. Furthermore, it is acknowledged that situational factors that are not related to trust also affect perceived reliance. What is however not acknowledged is how other factors may change over time, and how they affect one another not only through interaction, but also through other everyday life practices.

A limitation to many studies on trust in human-agent interaction and human-robot interaction so far is that they have mainly been performed in laboratory or experimental settings (e.g. Rossi et al., 2017; Agrawal and Yanco, 2018). This is understandable due to technology readiness levels to date. However, there are some studies that are starting to be performed in real-life settings, such as with senior citizens in care facilities and domestic environments (de Graaf et al., 2015; Klamer and Allouch, 2010; Tsiourti et al., 2014; Wada and Shibata, 2007; Wada et al., 2005, 2004; Broadbent et al., 2016). These studies mainly focus on agents as companions indicating positive effects on health and psychological well-being of people with respect to mood, loneliness and social connections with others. To our knowledge though, none of these studies focused on trust or collaboration per se, instead focussing mainly on acceptance aspects of the agents and derived implications for design.

4 Discussion

CSCW has a body of research on trust, focusing on computer-supported collaboration between humans. However, agents will increasingly become part of collaborative ensembles in the near future, such as in mixed human-robot teams in care contexts. Therefore, both taking into account agents as collaborative actors in different group formations, and exploring trust in agents in various collaborative settings, are important for understanding current and future collaboration.

As summarized in Table 1, we can broadly characterize some key differences in research on trust across different communities. For trust in CSCW when taking into account the broader context, trustees are mostly humans or institutions such as companies or governmental organizations, whereas in HAI and HRI, the trustee is the agent. A research gap for trust is the incorporation of agents as part of collaborative ensembles, where humans and institutions as well as agents can be trustees. For collaboration, trust in CSCW mostly refers to trust in human-human collaboration, often as part of larger ensembles, whereas in HAI and HRI, the focus is on human-agent collaboration, often as part of dyadic interactions. It is likely that different group formations with agents will have a different impact on trust, which is yet open for research. Regarding the study context and focus, trust in agents has so far largely been conducted in lab-based experiments asking questions around interactions, acceptance and so on, while CSCW research on trust has been conducted via field studies of everyday settings, trying to understand everyday practices and situated actions in complex contexts.

	Trust in Context	**Trust in H-A Interaction**
Trustee	Humans, Institutions	Agents
Collaboration	Human-Human	Human-Agent
Study Context	Field studies	Lab Experiments
Focus	Practices, Contextual Knowledge	Interactions, Acceptance

Table I: A broad characterisation of trust across different research communities.

As agents such as robots are likely to increasingly become part of collaborative ensembles in real-world settings, we argue that the CSCW field has much to contribute to HAI/HRI research, to take into account the complexity of dynamic environments for trust in agents. While lab-based studies can be fruitful for studying an interaction itself, we would gain important practical knowledge from taking contextual aspects and changes over time into account and how these relate to trust. For example, starting with the model proposed by Billings et al. (2012), we could make use of taking several aspects like environmental, person-related and agent-related factors into account and study how they dynamically affect trust. Furthermore, the three-factor model proposed by Hoff and Bashir (2015) involves temporal trajectories, where trust is dynamically co-shaped by interaction experiences. However, this approach has to our best knowledge not yet been applied to real-life settings with agents with regards to how everyday practices interrelate with *learned trust* and other factors that are part of the model. The many years of CSCW research studying collaborations over time and exploring notions of trajectories Fitzpatrick (2003) may have much to contribute here. We also argue that CSCW can broaden its concerns to also consider the role of agents in collaboration.

In support of our argument to include contextual knowledge in future research on trust in agents, theoretical work has pointed out that trust is not a "dyadic phenomenon between two isolated actors; there is usually always a context and a history, and there are also other actors that matter" (Möllering, 2006, p. 9). This relates to more recent HCI research, where practice-oriented approaches have been emerging (Kuutti and Bannon, 2014). Framing trust as it is enacted through ongoing practices as proposed by Corbett and Le Dantec (2018) or as a practice itself (Van House et al., 1998) is also a possible direction to go in to better understand the situatedness and dynamics of trust in collaborative ensembles that involve agents.

5 Conclusion

CSCW has a body of research on trust in computer-supported collaboration and relationships. However, agents will increasingly become part of collaborative ensembles in the near future, which should be taken into account in future research on collaboration. As trust is a key aspect of collaboration, trust in agents must be further explored. Despite research on the topic in other fields such as HAI and HRI, there is no agreed concept of trust in agents, and trust has been studied mainly as a psychological state or intention divorced from its context. Opposed to earlier interaction approaches, CSCW and HCI are strong in understanding the importance of collaboration in context, also embracing more practice-oriented approaches in recent years. In order to understand the complexitity, dymanics and situtatedness of trust, trust in agents may be better framed as interwoven with everyday practices, where CSCW can have a key role in conceptualizing and exploring how trust is part of collaborative ensembles that include agents.

References

Agrawal, S. and H. Yanco (2018): 'Feedback Methods in HRI: Studying Their Effect on Real-Time Trust and Operator Workload'. In: *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. New York, NY, USA, pp. 49–50, ACM.

Al-Ani, B., M. J. Bietz, Y. Wang, E. Trainer, B. Koehne, S. Marczak, D. Redmiles, and R. Prikladnicki (2013a): 'Globally Distributed System Developers: Their Trust Expectations and Processes'. In: *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*. New York, NY, USA, pp. 563–574, ACM.

Al-Ani, B., D. Redmiles, C. R. de Souza, R. Prikladnicki, S. Marczak, F. Lanubile, and F. Calefato (2013b): 'Trust in Virtual Teams: Theory and Tools'. In: *Proceedings of the 2013 Conference on Computer Supported Cooperative Work Companion*. New York, NY, USA, pp. 301–306, ACM.

Billings, D. R., K. E. Schaefer, J. Y. Chen, and P. A. Hancock (2012): 'Human-robot interaction: developing trust in robots'. In: *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction - HRI '12*. Boston, Massachusetts, USA, p. 109, ACM Press.

Billings, M. and L. A. Watts (2007): 'A safe space to vent: Conciliation and conflict in distributed teams'. *Springer, London*.

Bos, N., D. Gergle, J. S. Olson, and G. M. Olson (2001): 'Being There Versus Seeing There: Trust via Video'. In: *CHI '01 Extended Abstracts on Human Factors in Computing Systems*. New York, NY, USA, pp. 291–292, ACM.

Broadbent, E., N. Kerse, K. Peri, H. Robinson, C. Jayawardena, T. Kuo, C. Datta, R. Stafford, H. Butler, P. Jawalkar, M. Amor, B. Robins, and B. MacDonald (2016): 'Benefits and problems of health-care robots in aged care settings: A comparison trial'. *Australas J Ageing*, vol. 35, no. 1, pp. 23–29.

Brosnan, S. F. and F. B. M. de Waal (2014): 'Evolution of responses to (un)fairness'. *Science*, vol. 346, no. 6207, pp. 1251776.

Coeckelbergh, M. (2012): 'Can we trust robots?'. *Ethics and Information Technology*, vol. 14, no. 1, pp. 53–60.

Corbett, E. and C. A. Le Dantec (2018): 'Going the Distance: Trust Work for Citizen Participation'. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 312:1–312:13, ACM.

Crabtree, A., P. Tolmie, and W. Knight (2017): 'Repacking 'Privacy' for a Networked World'. *Springer, London*.

De Cindio, F., A. De Marco, and L. A. Ripamonti (2007): 'Enriching Community Networks by Supporting Deliberation'. *Springer London, Dordrecht Amsterdam*.

de Graaf, M. M. A., S. B. Allouch, and T. Klamer (2015): 'Sharing a life with Harvey: Exploring the acceptance of and relationship-building with a social robot'. *Computers in Human Behavior*, vol. 43, pp. 1–14.

Egger, F. N. (2000): '"Trust Me, I'm an Online Vendor": Towards a Model of Trust for e-Commerce System Design'. In: *CHI '00 Extended Abstracts on Human Factors in Computing Systems*. New York, NY, USA, pp. 101–102, ACM.

Fitzpatrick, G. (2003): *Locales Framework: Understanding and Designing for Wicked Problems*. Norwell, MA, USA: Kluwer Academic Publishers.

Garnik, I. (2004): 'Building Website Credibility: A Prospective Solution to e-Commerce in Poland'. In: *CHI '04 Extended Abstracts on Human Factors in Computing Systems*. New York, NY, USA, pp. 1019–1020, ACM.

Greenspan, S., D. Goldberg, D. Weimer, and A. Basso (2000): 'Interpersonal Trust and Common Ground in Electronically Mediated Communication'. In: *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*. New York, NY, USA, pp. 251–260, ACM.

Gulati, S., S. Sousa, and D. Lamas (2017): 'Modelling Trust: An Empirical Assessment'.

Gulati, S., S. Sousa, and D. Lamas (2018): 'Modelling Trust in human-like technologies'. In: *IndiaHCI '18*. pp. 40–61.

Hancock, P. A., D. R. Billings, K. E. Schaefer, J. Y. C. Chen, E. J. de Visser, and R. Parasuraman (2011): 'A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction'. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 53, no. 5, pp. 517–527.

Hoff, K. A. and M. Bashir (2015): 'Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust'. *Human Factors*, vol. 57, no. 3, pp. 407–434. PMID: 25875432.

Hong, J. (2009): 'Session Details: Privacy and Trust'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. –, ACM. Session Chair-Hong, Jason.

Human-Agent Interaction (2019): 'What is HAI? | Human-Agent Interaction'. http://hai-conference.net/what-is-hai [Online; accessed 7. Apr. 2019].

Jung, M. F., D. DiFranzo, B. Stoll, S. Shen, A. Lawrence, and H. Claure (2018): 'Robot Assisted Tower Construction - A Resource Distribution Task to Study Human-Robot Collaboration and Interaction with Groups of People'. *arXiv preprint arXiv:1812.09548*.

Kim, Y., A. Shaw, H. Zhang, and E. Gerber (2017): 'Understanding Trust Amid Delays in Crowdfunding'. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. New York, NY, USA, pp. 1982–1996, ACM.

Klamer, T. and S. B. Allouch (2010): *Acceptance and use of a social robot by elderly users in a domestic environment*. IEEE.

Knowles, B., M. Rouncefield, M. Harding, N. Davies, L. Blair, J. Hannon, J. Walden, and D. Wang (2015): 'Models and Patterns of Trust'. In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. New York, NY, USA, pp. 328–338, ACM.

Kostakos, V. and I. Oakley (2009): 'Designing Trustworthy Situated Services: An Implicit and Explicit Assessment of Locative Images-effect on Trust'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 329–332, ACM.

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*. New York, NY, USA, pp. 3543–3552, ACM.

Lampinen, A., V. Bellotti, C. Cheshire, and M. Gray (2016): 'CSCW and theSharing Economy: The Future of Platforms As Sites of Work Collaboration and Trust'. In: *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion*. New York, NY, USA, pp. 491–497, ACM.

Lee, J. D. and K. A. See (2004): 'Trust in Automation: Designing for Appropriate Reliance'. *Human Factors*, vol. 46, no. 1, pp. 50–80.

Lee, M. K. (2018): 'Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management'. *Big Data & Society*, vol. 5, no. 1, pp. 2053951718756684.

Linot, B. (2018): 'Trust in Computer-Supported crisis management information sharing'. *European Society for Socially Embedded Technologies (EUSSET)*.

Marsh, S. and J. Meech (2000): 'Trust in Design'. In: *CHI '00 Extended Abstracts on Human Factors in Computing Systems*. New York, NY, USA, pp. 45–46, ACM.

Martelaro, A., C. Nneji, A. Ju, and A. Hinds (2016): 'Tell me more: Designing HRI to encourage more trust, disclosure, and companionship'. In: *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. pp. 577–577.

Mayer, R. C., J. H. Davis, and F. D. Schoorman (1995): 'An Integrative Model of Organizational Trust'. *The Academy of Management Review*, vol. 20, no. 3, pp. 709.

Möllering, G. (2006): *Trust: Reason, Routine, Reflexivity*. Emerald Group Publishing Limited.

Nguyen, D. T. and J. Canny (2007): 'Multiview: Improving Trust in Group Video Conferencing Through Spatial Faithfulness'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 1465–1474, ACM.

Quan-Haase, A. and B. Wellman (2005): 'Local Virtuality in an Organization: Implications for Community of Practice'. *Springer London, Dordrecht Amsterdam*.

Riegelsberger, J. (2003): 'Interpersonal Cues and Consumer Trust in e-Commerce'. In: *CHI '03 Extended Abstracts on Human Factors in Computing Systems*. New York, NY, USA, pp. 674–675, ACM.

Robert, Jr., L. P. (2016): 'Monitoring and Trust in Virtual Teams'. In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. New York, NY, USA, pp. 245–259, ACM.

Robinette, P., W. Li, R. Allen, A. M. Howard, and A. R. Wagner (2016): 'Overtrust of Robots in Emergency Evacuation Scenarios'. In: *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*. Piscataway, NJ, USA, pp. 101–108, IEEE Press.

Rossi, A., K. Dautenhahn, K. L. Koay, and J. Saunders (2017): 'Investigating Human Perceptions of Trust in Robots for Safe HRI in Home Environments'. In: *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. New York, NY, USA, pp. 375–376, ACM.

Rousseau, D., S. Sitkin, R. Burt, and C. Camerer (1998): 'Not so different after all: A cross-discipline view of trust'. *Academy of Management Review*, vol. 23, no. 3, pp. 393–404.

Semaan, B. and G. Mark (2011): 'Creating a Context of Trust with ICTs: Restoring a Sense of Normalcy in the Environment'. In: *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*. New York, NY, USA, pp. 255–264, ACM.

Tsiourti, C., E. Joly, C. Wings, M. B. Moussa, and K. Wac (2014): 'Virtual Assistive Companions for Older Adults: Qualitative Field Study and Design Implications'. In: *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*. ICST, Brussels, Belgium, Belgium, pp. 57–64, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

Van House, N. A., M. H. Butler, and L. R. Schiff (1998): 'Cooperative Knowledge Work and Practices of Trust: Sharing Environmental Planning Data Sets'. In: *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work*. New York, NY, USA, pp. 335–343, ACM.

van Straten, C. L., J. Peter, R. Kühne, C. de Jong, and A. Barco (2018): 'Technological and Interpersonal Trust in Child-Robot Interaction: An Exploratory Study'. In: *Proceedings of the 6th International Conference on Human-Agent Interaction - HAI '18*. pp. 253–259, ACM Press.

Wada, K. . and T. . Shibata (2007): 'Living With Seal Robots - Its Sociopsychological and Physiological Influences on the Elderly at a Care House'. *Trans. Rob.*, vol. 23, no. 5, pp. 972–980.

Wada, K., T. Shibata, T. Saito, K. Sakamoto, and K. Tanie (2005): 'Psychological and Social Effects of One Year Robot Assisted Activity on Elderly People at a Health Service Facility for the Aged'. *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pp. 2785–2790.

Wada, K., T. Shibata, T. Saito, and K. Tanie (2004): 'Effects of robot-assisted activity for elderly people and nurses at a day service center'. *Proceedings of the IEEE*, vol. 92, pp. 1780–1788.

Wagner, A. R., P. Robinette, and A. Howard (2018): 'Modeling the Human-Robot Trust Phenomenon: A Conceptual Framework Based on Risk'. *ACM Trans. Interact. Intell. Syst.*, vol. 8, no. 4, pp. 26:1–26:24.

Wang, Y. and G. Mark (2013): 'Trust in Online News: Comparing Social Media and Official Media Use by Chinese Citizens'. In: *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*. New York, NY, USA, pp. 599–610, ACM.

Wang, Y. and D. Redmiles (2016): 'The Diffusion of Trust and Cooperation in Teams with Individuals' Variations on Baseline Trust'. In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. New York, NY, USA, pp. 303–318, ACM.

Waytz, A., J. Heafner, and N. Epley (2014): 'The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle'. *Journal of Experimental Social Psychology*, vol. 52, pp. 113–117.

Zheng, J., E. Veinott, N. Bos, J. S. Olson, and G. M. Olson (2002): 'Trust Without Touch: Jumpstarting Long-distance Trust with Initial Social Activities'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 141–146, ACM.

Konstantin Aal, Helmut Hauptmeier (2019): Pokémon GO: Collaboration and Information on the GO. In: Proceedings of the 17th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing and the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep04

Pokémon GO: Collaboration and Information on the GO

Konstantin Aal, Helmut Hauptmeier
University of Siegen
{firstname.lastname}@uni-siegen.de

Abstract. Pokémon GO is one of the few location-based mobile games (LBMG) which gained popularity all over the world. It increases physical activity when players are walking around catching Pokémon. It also shapes the players' sense of place and increases their social interactions. This exploratory auto-ethnographic study seeks to provide a first glimpse at how players appropriate different tools to inform themselves, collaborate with other players to catch or trade Pokémon and fulfill tasks inside the game together with a worldwide community. Results indicate that young adults learn how to organize catching events, arrange raid sessions, and collaborate within the Pokémon GO world and outside in the real world. This implies that the observed skills resulting from the gamification design elements of Pokémon GO can be transformed into the work life of young adults. Our explorative paper tries to pave the way for other research.

Introduction

Over the past decade smartphones have become ubiquitous, people use it all day long to inform themselves, to watch videos or play games. In addition, these devices have a huge number of sensors such as GPS, gyroscopes, and cameras which collect data about the usage of the smartphone and also about the users' surroundings. Based on these data, the smartphone can provide information about the environment such as restaurant recommendations or the shortest route to a location. These capabilities open many opportunities for developers to build

applications using the sensors to provide an immersive experience compared to traditional playing (in front of a computer or console). This leads to the creation of location-based mobile games such as “Ingress” (Chess, 2014) or “Insectopia” (Peitz, Saarenpää, & Björk, 2007). These games use the GPS-sensor of the smartphone to map real-world movements into the game: Players have to go to certain places where they can interact with the surroundings on their smartphone, for example, capture an arena, collect items or fight against other players.

One of the most successful location-based mobile games is Pokémon GO with over 65 million monthly active users where the majority of the players (78%) are between the ages of 18-34 (‘93 Amazing Pokemon Go Statistics’, 2016). Pokémon GO follows the tradition of the Pokémon anime and games, where players catch animal-like beings, the so-called ‘Pokémon’, and train them to battle others in the game. Pokémon GO goes one step further and lets players catch these Pokémon in real life on their smartphone or tablet. They can team up with other players and walk around in the real world where Pokémon will appear from time to time. The combination of a well-known and powerful brand (Pokémon) and the augmented reality (AR) experience of the game lead to the success story, and made even players who hadn’t played any mobile game in the past, start to catch Pokémon (‘Analysis of Pokémon GO’, n.d.).

Most of the previous studies focus on different aspects such as movement (Andone, Blaszkiewicz, Böhmer, & Markowetz, 2017), physical activity (Althoff, White, & Horvitz, 2016), engagement (Pyae, Luimula, & Smed, 2017), game mechanics (Tong et al., 2017), and social interactions (Paasovaara, Jarusriboonchai, & Olsson, 2017). Here we follow the traditional Computer-Supported Cooperative Work (CSCW) approach and observe how players appropriate new tools to gather information, collaborate with other players and acquire new skills.

Our research question tries to provide exploratory impressions about the issue of collaboration in a massive single player game with geolocation and multiplayer elements, in terms of 1) how do players especially young adults interact with each other to achieve their goals and 2) how can these collaboration patterns and skills be used and transferred in the work-life. These questions are not covered in detail in this paper, but provide first glimpses for future research.

Pokémon GO

In Pokémon GO players try to catch Pokémon, which will spawn when players start the smartphone app and walk in real-time through their neighborhoods or the outside world in general. One goal is to catch and collect a diversity of Pokémon. Based on the anime series, there are different fictional areas (e.g. Kanto, Hoenn, Johto) with different Pokémon. While walking through the real world, players will reach Pokéstops and visit Gyms in the Pokémon world. These locations are

virtual places where players can obtain items such as health potions, eggs, or Pokéballs about every five minutes or fight with other players or so-called raid bosses, i.e. a variety of Pokémon that control a Gym for a certain span of time. In addition, Players also have the opportunity to hatch eggs. These eggs are randomly collected at Pokéstops and hold rare or stronger Pokémon. They can be hatched by walking a fixed distance (2km, 5km, 7km, 10km).

One of the few multiplayer functions is the opportunity to fight other players at Gyms. Players can conquer these places and leave Pokémon inside the Gym. These Pokémon will defend the Gym against other players and the owner receives experience points and other rewards.

The latest updates added some new features to the game like trading Pokémon, battling other players in different leagues, or taking pictures of Pokémon and sharing them. The long-awaited feature to exchange Pokémon with others enables players to first become "friends" with other players. There is also a level-up system between this concept of friends. After the players accept the invitation to be friends, they can enter a virtual room to exchange Pokémon. To do so, they need to be in close distance to each other. Trading is also costly, and some trades can only be effected once per day.

Figure 1: (Left) Game environment, (Middle) Pokéstop, (Right) Gym with a Raid Boss Pokémon

Another important update was the Raid System. The Raid System extends the current Gyms by spawning very strong (and sometimes rare) Pokémon, called Raid Bosses, at these Gyms some of which can only be defeated when several players collaborate. The winners are awarded with the opportunity to catch the defeated Pokémon. A special version of these Raids is an EX-Raid. Players having attended several raids on one gym, have the chance to receive an invitation to an EX-Raid at a specific time and date. Only these players can fight against a very rare Pokémon and perhaps catch it after a victory. Raiding in groups is rather

weakly organized within the game as every player belongs to one of three teams (called Valor, Mystique, and Instinct, or Red, Blue, and Yellow). For raiding efficiently players can join "private groups" in order to increase their chances for catching the raid boss. Successful high-level raids inevitably require three or more players to be at the real-world place of the Gym. Although players are in fact informed about raid times and places in their vicinity, there is no guarantee that other players will be around and join in.

Niantic, the software development company of this game, also introduced so-called 'Community Days' (CD) to the Pokémon GO world. On a special day, the appearance of a special Pokémon increases. This event usually attracts very many players, who then walk through the city together and try to catch the special Pokémon (see Figure 2). Apart from CD the game also has 'special raid weeks', 'lunch break events' or 'Pokémon GO Fests'. Those fests attract more than 100.000 players per weekend and take place in cities like Yokusuka (Japan), Chicago (USA) or Dortmund (Germany). For 2019 an Earth Day Event will be organized to engage players in removing garbage from their real environment.

Successful raids or attacks at gyms require strategic thinking about what others, including the manufacturer, are thinking or will be doing. So metagame is an integral part of Pokémon GO and has important implications for the interactions between players.

Figure 2: Players during a Community Day.

State of the Art

In this study, we look at how (young) adults collect information and communicate with each other in communities; as another aspect, collaboration is generally relevant in games (such as online games, but also location based mobile games) and in particular in the game studied here, Pokémon GO.

(Young) Adults and Communication

Nowadays young adults regularly use a broad variety of mainstream communication tools – including social networking sites (Boyd, 2014), text, and instant messaging (Anandarajan, Zaman, Dai, & Arinze, 2010; Bouhnik & Deshen, 2014), cell phones (Alison Bryant, Sanders-Jackson, & Smallwood, 2006) and video calling (Buhler, Neustaedter, & Hillman, 2013) – in their everyday routines. Choi et al. found out that convenience plays a big role in the introduction of new instruments and technologies, convenience means above all availability at home, on the road and at work (David Choi, Chen, Wu, Lauterbach, & Balakrishnan, 2015). The study by Bouhnik et al. observed the challenges of integrating WhatsApp into classrooms. In particular, not all students could afford smartphones, the unpleasant interaction between students and the use of informal language were identified (Bouhnik & Deshen, 2014).

There is a long tradition in the CSCW research community to investigate how technology can be used to support collaboration in the workplace. Previous research has looked at the factors, that make the introduction of groupware unsuccessful, have discovered that the tools often do not fail due to technical problems, but due to a lack of understanding of the work environments that should support them (Grudin, 1994). Previous studies have shown that individual awareness and group activities are crucial for successful collaboration (Dourish & Bellotti, 1992). Bødker et al. took a closer look on an urban organic food community and examined the different artifacts that this community used to support their practice form; the authors describe it as ‘community artifact ecology’. This concept is helpful and important for CSCW, since it ‘enables framing of the between communities and technologies beyond the single artifact’ (Bødker, Korsgaard, & Saad-Sulonen, 2016).

Collaboration in Games

From the very beginning of video games, researchers were fascinated by the psychological and behavioral results of games and sought to understand the design characteristics responsible for the rich and varied motivational experiences and behavioral effects of games (Dongseong Choi & Kim, 2004; Yee, 2006). The first generation of LBMGs was mainly focused on serious applications, with a

focus on educational outcomes. The following generations have shifted the emphasis to playful attitudes and the social side of gaming (Hjorth & Richardson, 2014). This can be observed in particular in the context of online games that use the Internet to bring players together. In these cases, the collaboration seems to develop effortlessly between people who may not even have had previous connections (Cole & Griffiths, 2007; Yee, 2006). Here studies could show that the players especially enjoy the social interaction and cooperation in such games greatly (Yee, 2006). The study by Marker and Staiano (2015) showed that cooperative game aspects lead to higher engagement with the games or game systems than competitive approaches and lead to weight loss in overweight adolescents, but they also highlight that the aspects of cooperation and competition has to be examined in future gaming interventions (Marker & Staiano, 2015). In addition, work by Taylor (2008) highlight that users of online games are 'pushing back against simplistic notions of gaming as a form of passive media consumption' (Taylor, 2007, p. 122).

This brief insight should show that active cooperation and collaboration add value to games and that companies have a strong focus on providing players with a good gaming experience, including contact with other players.

Pokémon GO

The majority of literature dealing with Pokémon GO focuses on the influence of the game in terms of physical activity (Althoff et al., 2016), engagement (Pyae et al., 2017), game mechanics (Tong et al., 2017), and social interactions (Paasovaara et al., 2017). Recent literature also observed the sociability Pokémon GO and contrasted their results with another location based mobile game (LBMG) and emphasize "the importance of socialization's impact on making games a part of everyday life practices of the players." (Tokgöz & Polat, 2018, p. 9)

The study by Comunello and Mulargia tried to understand inter-generational interaction in the game and see Pokémon GO as "an environment that can potentially host meaningful interactions." (Comunello & Mulargia, 2017, p. 238) In a different study, the researchers analyzed data from mobile networks and present the effects of Pokémon GO on the "pulse of the city", which lead to more people being outside at different times and in people who were adapting their daily routines to the game play (Graells-Garrido, Ferres, Caro, & Bravo, 2017).

Pokémon GO affects the lives of many LBMG players on multiple levels. The social aspects are especially important, leading to more contact between players (which other games like Ingress have done as well) and beyond age limits. But so far only few studies have looked at the communication behavior and collaboration potential of the players and of the game itself. In particular the surrounding ecosystem of different tools (e.g. Blogs, Chat-groups, YouTube channels, maps,

discussion boards, messenger apps etc.) has not yet been investigated in any study.

Research Gap

We situate this explorative study between the communication behavior of young adults in their everyday life and collaborations in games and how these insights could be transformed into work life. Many studies have already addressed the collaboration aspect in online and offline gaming, but with the appearance of Pokémon GO, for the first time there is a LBMG which has reached a critical number of players (more than 5 million active players a day) and, in addition, has developed an ecosystem of tools used by the community (community artifact ecology).

Methods

For our exploratory study we used different methods to understand the research setting, the players (ourselves and other players) and the in-game collaboration patterns in-between players. An auto-ethnographic approach combined with qualitative methods should hence be appropriate to come to some first observations and insights into the relationships between gaming practices and its possible effects and implications on work life.

Auto-ethnographic Research

Auto-ethnography includes a reflexive and analytic account of personal experience and tries to connect this experience to wider social and cultural groups (Ellis & Bochner, 2000; Holman Jones, 2007). This method can be applied more rapidly to gain insights faster – the investigator inherits two roles: the informant 'insider' and the analyst 'outsider' (Cunningham & Jones, 2005).

The auto-ethnographic approach is justified since both authors are active players since the release of the game in July 2016, utilize different tools, and apply metagame strategies. The first author is an active member of several WhatsApp-Groups (three in total, with each of them for different purpose: Raids in two different cities and Pokémon GO task group), a Telegram channel (with general information about current developments in the game), follows Twitter users and also YouTube channels. The second author is also playing on a daily basis, but not active in any group or following someone on Twitter or YouTube.

Qualitative Methods

Besides observing how we interacted with other players, we also participated in chat groups and conducted informal conversations when attending raids or community days, therefore we adapted the Participatory Action Research (PAR) approach (Kemmis & McTaggart, 2005). Our results are based on empirical data collected from observations and informal conversations with actors participating in raids or other Pokémon GO events. Here, we were talking to and observing more than 30 players, with the age ranging from 14 to 69 years. About ten (age 18 to 28) of these players were tracked over a longer period of time online (WhatsApp) and offline, during community days and raid activities. All of them were aware that we use the provided information for research purposes. These empirical findings helped us better understand how the other players acted and appropriated different tools.

Data Analysis

Over the last six months we critically observed how we as players interacted with other players in the online and offline world. Therefore, we played on a daily basis, where the duration of the playing session differed each day depending on the given time, the group activities and special events. The majority of the time we were playing alone to catch and hatch Pokémon, but the first author was also monitoring the different WhatsApp-Groups to attend raid activities.

Field notes were written about these experiences every few days to capture the most important insights and screenshots of messenger chats as well as information channels to help us understand how information flowed from one channel to another. We then applied a thematic analysis to identify patterns of collaboration between Pokémon GO players. To achieve this, we first identified the different ways and tools, how players gather information and share with each other.

Afterwards, we tried to understand how these tools and the behavior of the players are intertwined with each other. We continue to monitor the group chats as well as blogs, Twitter, YouTube and social network postings of the individuals we encountered.

The collection of additional data (observations and informal conversations) confirmed the autoethnographic observations. It can be stated that autoethnography is a practical first step for understanding the activities and behaviors of users, enriched with additional empirical data these findings can provide a first glimpse into a new setting. The authors don't consider themselves as 'young adults', but use the same methods and tools as the younger players of Pokémon GO. Especially since the research focus of the authors circles around CSCW, this helps grind up the results correctly.

Findings

The following chapters describe first the different collaboration and information tools available to and used by players. Not all players use these tools, yet they obtain the same information in different ways.

Collaboration and Information Tools

WhatsApp. The first author was invited to participate in different WhatsApp-groups for his hometown. Pokémon GO players utilized the groups to inform each other about current developments of the game itself, about current or up-coming events such as Pokémon GO Community Days or strategies for raiding. Many players don't post much in the chats, but still use this opportunity to attend raid fights and get to know other players. While waiting for raid fights to start players usually discuss strategic or tactical aspects of how to approach the raid boss or they trade Pokémon.

Telegram. The chat app Telegram allows users to create channels, messages can only be spread across such channels. The Pokémon channel was used to receive the latest information about e.g. future events, updates and Pokémon sightings. The latter was particularly important for users who don't use Twitter or are not members of WhatsApp groups, since Telegram channels work without any invitation. Apart from that Telegram has features for polls and also access to Google maps for locating Gyms and Pokéstops.

Twitter. Players follow the official Twitter account of Niantic and many high-level players inform their followers about their daily activities or on how they achieved their high-level scores and on strategies for fighting in raids. In addition, Niantic's twitter account tweeted about the game and in-game developments and achievements (e.g. how many specific tasks, called 'research quests' were successfully finished).

YouTube. In the last two years, many Pokémon Go players started their own channels to broadcast their daily activities: hatching eggs, catching rare Pokémon, attending events or soloing high-level raid bosses. These players also speculate about future developments like subsequent Pokémon generations or the meta relevance of defense and attack Pokémon.

Forums and Blogs. Several forums and blogs focused on collecting and providing information around Pokémon Go. Many of these websites create infographics about raid fights (e.g. which Pokémon has the best attacks against certain raid Pokémon) and also rely on the players to gather information about the appearing spots of specific Pokémon (especially of rare Pokémon). Websites like 'pokemongohub.net' or 'pokemongo.gamepress' provide databases, wikis and tools for calculating internal values of Pokémon, their combat power or catch rate. These blogs focus on a broad variety of topics such as commercial interests of

Niantic, dissimilarities between rural and urban areas in terms of catching rates, availability of gyms, Pokéstops and other players. In some cases, these forums also present results from analyzing the code of every update of the game.

Collaboration Patterns

Existing Friends and In-game Socialization. The majority of players, but especially young adults, caught Pokémon with their real-world friends who were also playing the game. They used their usual chat application to organize their walks and meetings. This changed with the introduction of raids: players have to fight in larger groups against Raid Bosses. The line between offline and online friends blurred. Players knew more about their chat partners in the group chats (e.g. a father posted a picture of his newborn, others posted about power outages in the city) and met for the first time during the raid fights in real life. For community days, players made a walking plan to reach out to as many Pokéstops and gyms as possible in order to catch Pokémon.

Raid vs. EX-Raid. The introduction of raids increased the social element of Pokémon GO. Players had to cooperate in groups now to win against the Raid Boss. For this reason, many new chat groups were established by young adults for raid organization purposes only and invited the older players to these groups. Players posted pictures of Raid Bosses and asked for help, others asked for joining in the event and meeting at the raid location. EX-raids however needed a little bit more planning in advance. Players had to be active for a full week in several raids to receive an EX-raid invitation for some specific date, time and place. The exclusive participants of such EX-raid groups chat about and post screenshots of their EX-raid passes and organize meetings around the gym where the EX-raid will happen. As players receive invitations to different EX-raid locations, chat communication was sometimes quite confusing and produced meta-communication for clarification.

Gym Fights. Gym fights can be mastered solo, but as more in-game coins can be earned by taking and holding a gym for a longer time, the players started to coordinate. Players posted in group chats which gyms they had taken and asked other players not to fight them. However, you must be kicked out of a gym to receive the bonus coins. That's why the gyms were released later by the other players.

Trading Pokémon and Gifts. Another important update that strengthened the bond between the players was the ability to trade captured Pokémon with each other. Players first added their real-world friends as in-game friends. But through the activity in the group chats, lists of player IDs were created to make it easier to make new friends in-game. It is also important to maintain these in-game friendships so that you can get different bonuses. In-game friendship also led to forms of organizing when to open or send gifts to each other. These presents

contain items and the more often they are exchanged the higher the level of friendship gets which in turn leads to increased chances of catching Raid bosses.

Exchange of Information. Participants who subscribed to YouTube channels shared their knowledge of how to best catch Pokémon. Twitter users, who were usually below 25 years old in our study group, shared screenshots of the official Pokémon Twitter account. Infographics shared on Telegram were also uploaded to the WhatsApp groups. WhatsApp was the central organizational tool where all information was collected. Due to the large size of the groups, there was always someone who could answer questions quickly. Also, tips were exchanged about tools that help catch Pokémon (although a few are not officially allowed). Here especially the young adults were using all of these tools and other social networking sites besides WhatsApp.

Going beyond Pokémon. Younger adults, who are more experienced using a smartphone helped older players in troubleshooting with their smartphone during raid fights or community days, even without knowing each other for a long time. Connecting with new people was an important aspect for many players.

Discussion

The discussion aims to highlight two aspects that might be relevant for work life: Collaboration and transferability of IN-game skills to the working environment. These are just some first tiny aspects which need further investigation, but might play a major role for young adults who just transferred into work life.

Pokémon GO and Collaboration

Collaboration was an important aspect in the game to successfully catch all Pokémon, as well as reach higher levels. It could be observed how experienced players took the new players under their wing and invited them into the groups. This is about the dissemination of information and also about forming effective strategies for advancing in the game. At the beginning these new players were mostly quiet and not very active in the chats. Through the social aspect (Tokgöz & Polat, 2018) of Pokémon GO, there was a lot of real-life exchange during the individual raid fights at the gym. Here the players were supposed to talk to each other so that there was no confusion and everyone had the chance to participate in the raid. Less experienced players are often given advice when it comes to selecting the optimal set of Pokémon for a fight or when investments of resources into the capabilities of certain Pokémon are at stake. At the same time, the group chats were also used to keep players who weren't there yet up to date, this collaborative behavior was already apparent in early online games (Yee, 2006).

Other collaboration opportunities were seen during the capture of gyms. Players formed smaller groups to occupy and defend multiple gyms in order to receive the bonus. These were mainly groups of players who already knew each other from the real world, who were then spontaneously enlarged by other, players, who were also active in the chat groups.

Pokémon GO and Work Life

Many of the skills acquired during the game play of Pokémon GO can be transferred into the work life such as *gathering information, strategic capabilities, social aspects, communication* and *flexibility*.

Information plays a major role in the working world, but also in the Pokémon GO world. Players use several ways to collect information (such as the location of a rare Pokémon, which tools are available and which Pokémon helps against certain opponents) and then bundle and process it. By providing this processed information to other players, they could strengthen the whole team. The information is quite complex in nature here: there are hundreds of different Pokémon, each belongs to one or two types out of 18 different types. Each Pokémon has internal values (attack, defense, stamina), two different types of attacks out of more than one hundred attacks in total etc. Attacks take different spans of time for loading and unloading etc. The properties of the Pokémon are interrelated and partly even dynamic (random attack boosts, influence of the real-world weather etc.). So, some of the Pokémon's behavior can be predicted from knowledge about its properties and some cannot because of randomness. And in between determinacy and randomness there are some behavioral aspects of Pokémon which are probabilistic. Experienced players know about the probability distribution of such behavior and instruct younger players about how to increase e.g. the probability of catching certain Pokémon. In a nutshell, information in Pokémon GO is quite dense, partially certain, highly interconnected and transfers apparently best when being demonstrated in situ (e.g. there is a catching trick one can read about on websites or watch on YouTube, yet a real-life demonstration seems to be more effective for learning it).

Another aspect is social skills: integrating into an existing team, making new friends and contributing to the overall goal. This could be seen at community days and raid fights, when several players, that had never seen each other before, formed groups. Pokémon GO players form a highly diverse set of people, with different lifestyles, differences in age, gender, occupation or life-worlds ("Lebenswelten") so-to-speak. So communicative skills are needed for addressing shared goals or getting information for one's own goals. This includes the ability to quickly form a team to perform certain tasks successfully, which is also relevant to both areas, Pokémon GO and the world of work, where an event often requires a quick response. This aspect is especially important in today's

increasingly complex and flexible work environment: employees need to be reachable through multiple communication channels (e.g. email, but also messenger and social media). By being part of a community artifact ecology (Bødker et al., 2016) players experience the overlapping ecologies of using different tools, discussing topics in related communities and observing the evolution of these communities.

Limitations and Future Work

This auto-ethnographic should provide a first glimpse at how collaboration between players and especially young adults in the realm of Pokémon GO happens. Since only the findings and experiences of two players and their observations of the real-world activities and chats are analyzed, this study is neither representative nor systematic for all players and for all their different goals and motives.

Future research should include more players with differences in age, educational, social and economic background. Especially young adults who just started their work life and are Pokémon players as well. In particular, qualitative methods such as semi-structured interviews and participatory observations should be used to gather meaningful insights. While we tried to focus on young adults, the majority of the results are based on the observations and talks with all player.

Conclusion

Our study wanted to show how Pokémon GO established ways of collaboration between players which can be transformed into work place contexts. Especially, reacting to time critical events (e.g. raid fights) or planned events (e.g. EX-Raids), where players gather rather quickly for fighting a Pokémon which would be too strong to be defeated alone. In addition, the players are building private groups during the raid process according to the team memberships and are helping each other afterwards to catch the raid boss.

Young adults' appropriate different tools and platforms to gather information, distribute them to other players and organize events to catch rare Pokémon. Especially older users are not familiar with many of these tools and platforms and profited from the skills of the younger ones. These skills are essential in a fast moving and flexible work environment and therefore helpful for the company if they could be integrated in the company's communication and organization infrastructure.

References

93 Amazing Pokemon Go Statistics. (2016, August 27). Retrieved 11 September 2018, from DMR website: https://expandedramblings.com/index.php/pokemon-go-statistics/

Alison Bryant, J., Sanders-Jackson, A., & Smallwood, A. M. (2006). IMing, text messaging, and adolescent social networks. Journal of Computer-Mediated Communication, 11(2), 577–592.

Althoff, T., White, R. W., & Horvitz, E. (2016). Influence of Pokémon Go on Physical Activity: Study and Implications. Journal of Medical Internet Research, 18(12), e315. https://doi.org/10.2196/jmir.6759

Analysis of Pokémon GO: A Success Two Decades in the Making. (n.d.). Retrieved 11 September 2018, from Newzoo website: https://newzoo.com/insights/articles/analysis-pokemon-go/

Anandarajan, M., Zaman, M., Dai, Q., & Arinze, B. (2010). Generation Y Adoption of Instant Messaging: An Examination of the Impact of Social Usefulness and Media Richness on Use Richness. IEEE Transactions on Professional Communication, 53(2), 132–143. https://doi.org/10.1109/TPC.2010.2046082

Andone, I., Blaszkiewicz, K., Böhmer, M., & Markowetz, A. (2017). Impact of location-based games on phone usage and movement: a case study on Pokémon GO. Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '17, 1–8. https://doi.org/10.1145/3098279.3122145

Bødker, S., Korsgaard, H., & Saad-Sulonen, J. (2016). `A Farmer, a Place and at least 20 Members- The Development of Artifact Ecologies in Volunteer-based Communities. Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16, 1140–1154. https://doi.org/10.1145/2818048.2820029

Bouhnik, D., & Deshen, M. (2014). WhatsApp goes to school: Mobile instant messaging between teachers and students. Journal of Information Technology Education: Research, 13(1), 217–231.

Boyd, D. (2014). It's complicated: The social lives of networked teens. Yale University Press.

Buhler, T., Neustaedter, C., & Hillman, S. (2013). How and why teenagers use video chat. Proceedings of the 2013 Conference on Computer Supported Cooperative Work - CSCW '13, 759. https://doi.org/10.1145/2441776.2441861

Chess, S. (2014). Augmented regionalism: Ingress as geomediated gaming narrative. Information, Communication & Society, 17(9), 1105–1117. https://doi.org/10.1080/1369118X.2014.881903

Choi, David, Chen, J., Wu, S., Lauterbach, D., & Balakrishnan, A. (2015). From Dorms to Cubicles: How Recent Graduates Communicate. 2015 48th Hawaii International Conference on System Sciences, 2013–2022. https://doi.org/10.1109/HICSS.2015.241

Choi, Dongseong, & Kim, J. (2004). Why people continue to play online games: In search of critical design factors to increase customer loyalty to online contents. CyberPsychology & Behavior, 7(1), 11–24.

Cole, H., & Griffiths, M. D. (2007). Social interactions in massively multiplayer online role-playing gamers. Cyberpsychology & Behavior, 10(4), 575–583.

Comunello, F., & Mulargia, S. (2017). My Grandpa and I "Gotta Catch 'Em All." A Research Design on Intergenerational Gaming Focusing on Pokémon Go. In J. Zhou & G. Salvendy (Eds.), Human Aspects of IT for the Aged Population. Applications, Services and Contexts (Vol. 10298, pp. 228–241). https://doi.org/10.1007/978-3-319-58536-9_19

Cunningham, S. J., & Jones, M. (2005). Autoethnography: a tool for practice and education. Proceedings of the 6th ACM SIGCHI New Zealand Chapter's International Conference on Computer-Human Interaction Making CHI Natural - CHINZ '05, 1–8. https://doi.org/10.1145/1073943.1073944

Dourish, P., & Bellotti, V. (1992). Awareness and coordination in shared workspaces. Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work - CSCW '92, 107–114. https://doi.org/10.1145/143457.143468

Ellis, C., & Bochner, A. P. (2000). Autoethnography, personal narrative, reflexivity: Researcher as subject.

Graells-Garrido, E., Ferres, L., Caro, D., & Bravo, L. (2017). The effect of Pokémon Go on the pulse of the city: a natural experiment. EPJ Data Science, 6(1), 23.

Grudin, J. (1994). Groupware and social dynamics: eight challenges for developers. Communications of the ACM, 37(1), 92–105. https://doi.org/10.1145/175222.175230

Hjorth, L., & Richardson, I. (2014). Gaming in social, locative and mobile media. Springer.

Holman Jones, S. (2007). Autoethnography. The Blackwell Encyclopedia of Sociology.

Kemmis, S., & McTaggart, R. (2005). Participatory Action Research: Communicative Action and the Public Sphere. Sage Publications Ltd.

Marker, A. M., & Staiano, A. E. (2015). Better together: outcomes of cooperation versus competition in social exergaming. Games for Health Journal, 4(1), 25–30.

Paasovaara, S., Jarusriboonchai, P., & Olsson, T. (2017). Understanding collocated social interaction between Pokémon GO players. Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia - MUM '17, 151–163. https://doi.org/10.1145/3152832.3152854

Peitz, J., Saarenpää, H., & Björk, S. (2007). Insectopia: exploring pervasive games through technology already pervasively available. Proceedings of the International Conference on Advances in Computer Entertainment Technology - ACE '07, 107. https://doi.org/10.1145/1255047.1255069

Pyae, A., Luimula, M., & Smed, J. (2017). Investigating Players' Engagement, Immersion, and Experiences in Playing Pokémon Go. Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition - C&C '17, 247–251. https://doi.org/10.1145/3059454.3078859

Taylor, T. (2007). Pushing the borders: Player participation and game culture. Structures of Participation in Digital Culture, 112–131.

Tokgöz, C., & Polat, B. (2018). Sociability on Location Based Mobile Games: An Ethnographic Research on Pokémon Go and Ingress in Istanbul. European Journal of Social Sciences Education and Research, 12(1), 120–129.

Tong, X., Gupta, A., Lo, H., Choo, A., Gromala, D., & Shaw, C. D. (2017). Chasing Lovely Monsters in the Wild, Exploring Players' Motivation and Play Patterns of Pokémon Go: Go, Gone or Go Away? Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17 Companion, 327–330. https://doi.org/10.1145/3022198.3026331

Yee, N. (2006). Motivations for play in online games. CyberPsychology & Behavior, 9(6), 772–775.

Ciolfi, L.; Lockley, E. (2019): Exploring Flash Fiction for the Collaborative Interpretation of Qualitative Data. In: Proceedings of the 16th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing an the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep03

Exploring Flash Fiction for the Collaborative Interpretation of Qualitative Data

Luigina Ciolfi, Eleanor Lockley
Sheffield Hallam University, UK
[l.ciolfi] [e.lockley]@shu.ac.uk

Abstract. This paper presents some exploratory reflections on flash fiction as a possible method to spark discussion and collaborative interpretation of qualitative research data. A growing body of work in HCI and CSCW examines the potential of techniques used in creative writing and creative fiction to generate design concepts, and narrative data analysis is adopted by social science using creative writing techniques for qualitative data work. Here we discuss our experience of an exercise where flash fiction was used not as a technique in support of design (which has been done before in human-centred computing), but as a means of probing data and facilitating collaborative data work among researchers. We reflect on the experience and outcomes of the exercise and also discuss exploratory ideas regarding how creative writing techniques could be further explored in human-centred computing as a way to probe findings from empirical data, particularly for collaborative teams.

Introduction and Related Work

Methodological innovation in CSCW and HCI is very often concentrating on the design phase of the technology development process, and the process of design itself is usually the focus of experimentation with research methods (usually adapted from other design disciplines and from creative practice fields, such as the performing arts) to spark creative ideation (Wright and McCarthy, 2004). In

between the requirements gathering phase and the design phase sit techniques such as Cultural Probes (Gaver, Dunne and Pacenti, 1999), which have the goal to gather input to inspire design and are a design method in itself, not a user research methodology as such.

On the other hand, methods for collecting, analyzing and mining empirical data tend to be much more traditional and slow to change: whether field ethnographies, or controlled experimental studies, or surveys or questionnaires, the techniques that CSCW and HCI researchers use for data gathering are very consolidated ones, and so are those for analyzing and making sense of the data (i.e. statistical analysis and/or qualitative data analysis), with little in the way of innovation and experimentation. This, of course, makes sense in light of the pursuit of scholarly rigor where ensuring the validity of a study is a priority to satisfy research funding terms and peer-reviewers, and a core necessity in order to base the design of systems on strong empirical evidence and clearly identified requirements. However, at the same time, as a community we are perhaps missing out on opportunities for methodological innovation beyond the focus on design, and relating to more on how we make sense, internalize and respond to data. As the goal of CSCW and HCI research is usually that of impacting on digital systems (whether by designing them, evaluating them, or investigating how they shape people's practices), there is space to explore additional and/or alternative ways to reflect on user data and in doing so providing additional bridges between data and requirements collection and creative ideation.

Methodological innovations in user research in CSCW are usually tied to systemic shifts in the discipline, for example the argument in favour of field research in the early 1990s that introduced ethnographic approaches into a community that had predominantly worked in (experimental) labs up to that point (Bødker, 20015). As part of these lines of disciplinary growth, another example is how techniques with strong storytelling focus (such as scenarios and personas) have become established part of envisioning methodologies leading to design concepts and prototypes (Carroll, 1995). Scenarios are narrative descriptions of intended user interactions and experiences that have not yet been realized, depicting also the situation of use, features of people involved, etc. They are used to both communicate ideas within a team and brainstorm around how digital interactions could take place in a particular setting, or with the involvement of different people, before any prototyping or testing is done (Carroll, 1995). Personas are rich descriptions of realistic –albeit fictional- characters intended to embody various characteristics of expected real-world users (e.g. different age, role, technical skills, nationality, etc.) so to represent various intended user groups in less abstract terms (Grudin, 2003). Personas have often been critiqued for they can reproduce assumptions or even stereotypes associated to a user groups (Putnam et al, 2009), and can represent and reinforce the "othering" of a certain community (see Cabrero, Winschiers-Teophilus and Abdelnour-Nocera, 2016).

These techniques are grounded in elements of storytelling, however they do not lead to fully fictional outcomes. Rather, they blend realistic elements based on actual situations of technology use or actual user groups with some imagined characteristics, which tend to be all the same realistic and achievable.

In relation to design, several fiction techniques have been used to generate fully speculative outcomes. Notable examples are Critical Design Fiction and Speculative Design Fiction (Dunne and Raby 2013; Auger 2013; Brown et al 2016; Comber et al 2018), which have been widely adopted in interaction design. In this case, the goal is not to represent a near-future, realistic scenario of use, but rather imaginary and provocative situations, which can be used to probe the impact and role of digital technologies beyond the realm of what is considered possible and beyond the focus on providing a solution to a problem (Blythe, 2014).

Care and attention towards the narratives of data and the narratives of design elaborations signal the key importance of stories, plots and characterization in the HCI and CSCW field. Blythe (2014) argues how storytelling and narrative approaches have not very often embraced fiction creation: as Blythe notes, scenarios are usually based on findings pertaining to real people and on other factual bases. More recently, several HCI researchers have explored the potential of methods and techniques commonly used in creative fiction for imagining and discussing novel design and interaction concepts. A notable example is design fiction (Sterling, 2009; Linehan et al. 2014; DiSalvo 2012; Blythe and Wright 2006; Smit et al. 2018), which has been examined as an approach to explore both technical and interactional aspects of prototypes, and issues and open questions around technology use (Blythe, 2014; Hales, 2013; Linehan et al., 2014). According to Blythe, fiction stands on a different plane compared to other narrative techniques such as scenarios:

> Television and film scripts are also usually written in the present tense. Other forms of fiction, like novels and short stories, seldom are, so why this difference? Perhaps because, scenarios, like scripts, are in a process of becoming: they are there to be made into something else. A scenario is part of a process, a fiction exists in its own right. Perhaps for this reason also scenarios do not end, rather they stop, they are not resolved. For resolution to occur a conflict must be worked out and this is another structural difference between scenarios and fiction.
>
> (Blythe 2014, p 4)

In more recent years, there have been attempts to develop and apply "hybridized" techniques blending scenarios and fictions, mainly as a way to question design values (see, for example, Muller and Liao 2017). Advocating a role for fiction is a provocative position to hold, in a space where there is constant preoccupation with process, impact and translation into functionality, and indeed

the appropriateness of addressing some research domain through fiction approaches has been questioned (see Iivari and Kuutti, 2018). However, it is an interesting topic to probe and explore, especially in that "in between" space from data analysis to design.

As we mentioned, data work usually occurs on the basis of qualitative or quantitative paradigms for collection and analysis. Particularly when it comes to qualitative research, we argue that there is value in establishing a relationship with data beyond an analytical one, in empathizing with the perspective of other people while generating new ideas and thoughts of the researcher's own, and in engaging with the imaginary. Therefore, we believe that the potential of using fiction methods for aspects of the human-centred design process other than creative design needs to be further studied.

In HCI and CSCW, narrative- and fiction-based techniques have indeed been used to engage informants and design participants in developing imagined futures and in considering aspect of technology that might not be immediately apparent or might benefit for more open-ended treatment, such as ethical and value implications (Cheon and Su, 2017). Short fictions have been used as prompts for probing a focus group (Draper and Sorell, 2014), and groups of external participants have been engaged by researchers to create participatory design fictions (Muller and Liao, 2017).

In our case, however, we are interested in looking at fiction techniques – particularly literary fiction – as tools not for end-users or informants to engage in, but for researchers themselves, and particularly those who deal with qualitative data.

The parallels between finding narratives in data and weaving imaginary narratives from inspiration have been explored in methodologies such as narrative research analysis, where the researcher is reconstructing narratives from "messy" data (Kim 2016). Narrative inquiry instead focuses on storytelling on a factual basis (research data itself), but the characterization of the researcher as "storyteller" rather than detached analytical voice is a powerful methodological positioning, as noted also by Wright and McCarthy (2004). Richardson (1994) argues that writing is important as a *method of inquiry* as well as a *method of knowing*. Furthermore, fictional accounts can lead to empathy:

> Qualitative researchers have come to believe that fictional accounts can sometimes portray a research phenomenon more clearly than do the standard representations of qualitative data (…) The fictionalisation of research data provides researchers with the opportunity to work with raw data in order to speak to the heart of the reader's social consciousness, while providing the protection of anonymity to the research participants.
>
> (Kim, 2016, p. 140)

We believe that such approaches can aid CSCW and HCI researchers tackle the long-existing challenge of moving from data analysis to "implications for design". Naturally, we do not argue that established techniques for data analysis should be ignored or replaced, but rather that additional techniques can be used to facilitate other aspects of working with data, particularly in collaborative teams designing together, and that they are worth exploring and discussing within our field.

Inspired to explore this methodological approach both by the body of work we have now briefly discussed, and also by our interests in how to communicate ideas and encourage reflection in collaborative, data-focused workshops, we designed and conducted an exercise exploring fiction writing as a technique for the collaborative reflection over qualitative data among researchers, which we present in the following section.

The Flash Fiction Workshop

The exercise took place as part of a research project that explored human practices of blurring and/or balancing work and life demands and of using digital technologies for these purposes (Ciolfi and Lockley 2018). After having conducted an interview study with 26 participants collecting empirical data on such practices, we disseminated a summary of our results as part of a one-day workshop called "Managing Technology Around Work and Life" aimed at other researchers interested in these topics, and involved the workshop participants in a creative exercise that adopted the technique of *flash fiction*, commonly used in creative writing.

Flash fiction (Galef, 2016) is an approach to creative writing where the author responds to one or more short prompts by creating a brief story in a short time frame, usually within one hour. Flash fiction generates writing that is longer than micro-fiction (which is usually less then 300 words) and much shorter than literary short stories of several thousand words. In addition, flash fiction has the characteristic of being less polished and rather a way to put ideas, reactions and explorations down on paper. Flash fiction is seen as both a way to maintain writer's creativity and flow of ideas, and to generate fiction that captures illuminating thoughts and reactions despite its short length and short preparation time.

Creative writers experimenting with flash fiction usually rely on *prompts*. Flash fiction prompts can take a variety of forms, from specific instructions or directions (e.g. "Write about a lonely child who finds a friend"), to out-of- context sentences or phrases to be extended and elaborated ("He was such a lonely child...He couldn't believe it when he realised that he had found a friend").

Flash fiction writing aids exist such as flash fiction notebooks and diaries with daily prompts, and toolkits containing prompts, drafting cards and other inspirational materials. A number of websites, mailing lists and online groups also offer daily prompts to subscribers. Examples are 3AM Magazine (http://www.3ammagazine.com/3am/), Everyday Fiction (http://everydayfiction.com/), Nano Fiction (http://nanofiction.org/category/weekly-feature/writing-prompts) and Brevity (http://brevitymag.com/).

Beyond creative writing, flash fiction as a technique has been used in education, for example to encourage students to write opinion pieces (Setyowati, 2016). It is also a popular technique in the digital fiction world, especially for collaborative fiction using microblogging platforms such as Tumblr and Twitter (Bell, Hesslin and Rustad, 2014; Shapard, 2012).

For our work/life project's workshop, we designed the prompts on the basis of the empirical data that had been collected in our interview study (for full details about the study see (Ciolfi and Lockley 2018)). We now very briefly describe the study to provide an idea of the type of data we gathered and its themes.

The interview study involved a sample of 26 people of working age (the youngest participant was 24 and the oldest 62) in knowledge-intensive roles in high employment sectors in the British city of Sheffield (education, IT, creative industries, design and engineering). 12 participants were women and 14 were men. Occupations included: Education/training consultant, Business Development Manager, Senior Producer, CEO, Information Officer, Strategic Development Manager, Knowledge Transfer Researcher, Designer, Librarian, Lecturer. The interviews were semi-structured, and participants were asked questions about themselves (educational background, professional role, etc.), the work that they do, some aspects of their private life, and about how they deal with the challenges and demands of work and life. They were also asked about their use of digital technology for managing their time and multiple demands. The interviews were audio-recorded and lasted between 40 and 90 minutes. The transcriptions were then analysed by the authors of this paper through repeated readings and the identification of thematic codes. The study captured a set of lived practices around work, life and the role of technology and the interviews provided detailed insights of the participants' perceptions, decisions and strategies.

In the workshop, excerpts from the interview data were used as flash fiction prompts. We selected and slightly edited (e.g. removing pauses and repetitions) the excerpts to make them work as prompts that would mimic those used in creative flash fiction.

Our rationale was to select quotes from different interviewees that could be thought-provoking, hold multiple meanings, or lead to multiple interpretations, to see which aspect the workshop participants would choose to focus on and develop, and in what way.

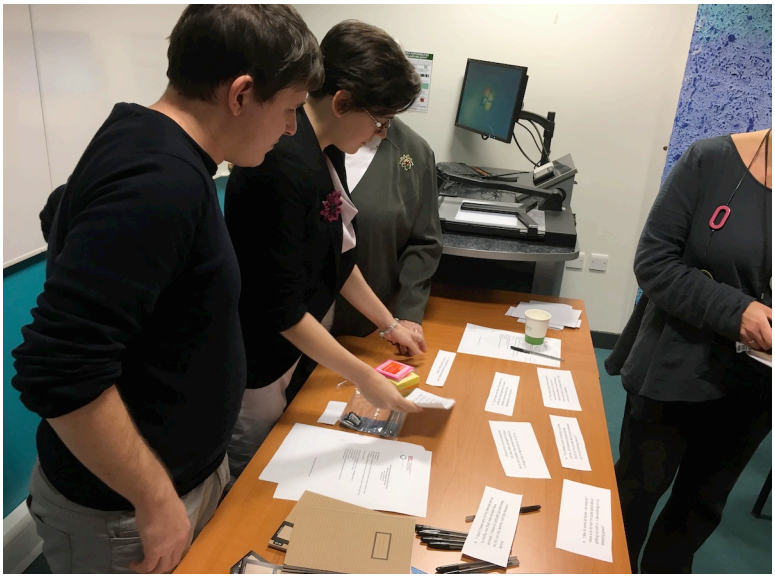


Figure 1. Workshop participants selecting flash fiction prompts for their story.

We chose to present the prompts without giving any information relating to the interviews they came from (i.e. what kind of person they involved, working in which sector, or speaking in relation to which context). This decision was motivated by the need to maintain the conditions of anonymization of the data, and in part to do with not wanting to tie the flash fiction exercise to the real-life circumstances of the interview respondents and therefore to scenarios that might be too closely related to them.

Twelve people took part in the event and they were a mix of academic staff and postgraduate students interested in the research topic of work/life boundaries and coming from the disciplinary areas of HCI and social science (cultural studies, communication and sociology). Some of the participants were colleagues from our university while others were external attendees from other institutions who had joined the workshop for the day.

All of them had experience of gathering and analyzing qualitative data and they were informed that the prompts for the flash fiction exercise were extracted from interview transcripts. The participants were not part of the original project and were recruited through academic mailing lists. None of them had any relationship to the interviewees from whom we had collected data.

The workshop started with short presentations (followed by a Q&A) by the authors of this paper about our project and about the interview study and its main thematic findings. The rest of the workshop was then dedicated to the flash fiction exercise, for which we took on the role of facilitators.

Four groups of three people were formed for the exercise, and each group was assigned an initial prompt by us. The motivation for this was to reduce the start-up time for the exercise and to distribute prompts that could lead to very different narratives.

The groups were briefed on flash fiction and on the modality of the exercise. The brief stated that each story had to feature interaction with technology that is not limited to what is technically possible or already existing. Participants could be as creative or speculative as they wished.

The groups were given two hours overall to produce their stories. After one hour, the groups were asked to choose a second prompt from a selection (Fig 1). They could either choose a prompt that pushed their story in a different direction, or provided a 'twist', or a prompt that supported the storyline that they had developed up to that point.

Figure 2. Plotting the flash fiction through post-its and simple storyboarding.

Examples of the 18 prompts we generated were:

- *"We've been in a variety of deserted desert islands with no electricity yet still been working, which is not ideal but these things do chase you around"*
- *"I just made sure that in that week I got a little bit of work done and I've sent emails to give the kind of the appearance of doing work"*
- *"Sometimes I physically feel like I want to and I have to stop myself"*

- *"I don't have any of my devices set up to notify me that new emails have arrived. I have to actually go and check, so it's on my terms, not the device's terms"*
- *"If I'm on holiday and I am not gonna look at anything work-related, then I'll pay the penance on the other side"*

The groups were given notebooks, sketching paper, pens and post-it notes to help them discuss, plot and organise their story (Fig 2; Fig 3). Two groups wrote their story by hand in the copybooks provided, two decided to write it using a laptop and word processor.

In responding to prompt 1 (*I need people to think my business is bigger than a one man band so I never have my office hours on my signature*), Group 1 devised a surreal and fantasy-laden story titled "Armorgeddon: There is a rhino loose in the city". "Armorgeddon" took aspects of work/life demands and blurring to extreme and thought-provoking paradoxes:

> "*He* [The Protagonist] *opened an online shop, Armorgeddon, selling a variety of weird and wonderful sea shells from around the world. He was adamant to run his shell business alone, but had big ambitions - he wanted the world to view his business as a considerably bigger entity than the reality. "I need people to think my business is bigger than a one-man band so I never have my office hours on my signature", he would think. One consequence of his approach was that he needed to spend long hours working, and always had to be on call – constantly checking his emails on his phone, and dealing with orders on his laptop.*
>
> *He situated his business and his life in a beautiful disused ivory tower looking over the local town's square, Shellington. He surrounded himself with Minions to help with his dastardly deeds… unfortunately his Minions were not real, and were instead faces painted onto balloons hanged against the windows of his tower, to give the correct impression. To anyone looking from the town square, it would appear he had an army of minions working for him*" (excerpt from "Armorgeddon")

Group 2 responded to prompt 2 (*If I don't focus on work when I'm at work, I could kill someone!*) with "Under Pressure", a story set in 2019 and about the crew of a deep-sea nuclear submarine stuck under the Polar ice cap in the dead of Winter. As the ice is too thick for them to surface, they cannot move and have limited provisions on board. The boundaries between life and work completely dissolve, although the professional roles that some of the crew members have (for example, the medical officer) make them decision-makers for issues to do, for example, with rationing food. Group 2 also selected background music to accompany the recitation of their story – the song "Under Pressure" by Queen and David Bowie.

Group 3 (Fig. 3) worked from the prompt *"We've been in a variety of deserted desert islands with no electricity yet still been working, which is not ideal but these things do chase you around"*.

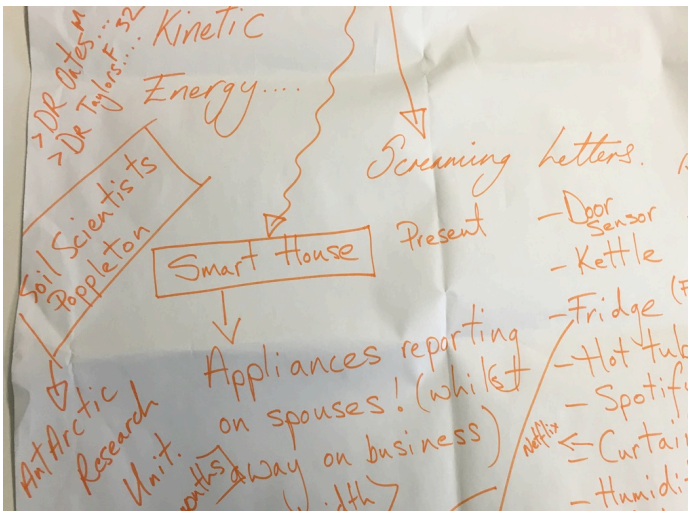


Figure 3. Group 3 plotting the story through brainstorming and keywords.

Their story is written as a set of log entries and automated notifications from smart home systems being received by academics currently on a field mission in Antarctica. It plays on how automated life-related interruptions in a physically remote and "extreme" work setting might be interpreted and dealt with in a context that is as far removed as possible from the location where they are received:

> *"[Wild Oates] 24 Railway Cuttings: 13/12/2016: 4.32PM: Spotify. Playlist. Kanye West.*
>
> *[Taylor Towers] 32 Windsor Gardens: 13/12/2016: 7.14PM: Curtains. Closed.*
>
> *[Taylor Towers] 32 Windsor Gardens: 13/12/2016: 8.16PM: Fridge. Waitrose shopping list ordered.*
>
> *Mission log. Prof. Pankhurst. Expedition Day 95. Getting tired of eating these Waitrose Essentials Ships' Biscuits. Dr Taylor has been complaining again about the limited bandwidth. I've already told him that Skype and Youtube are not essentials. However, we do frustratingly seem to have the bandwidth for Oates and Taylor to both keep receiving their smart home updates. Personally, I have to turn my phone off at night or it'll constantly vibrate with notifications and email alerts. Vodka has still not arrived.*
>
> *[Taylor Towers] 32 Windsor Gardens: 14/12/2016: 8.00AM: Fridge. Waitrose Fishy Friday Oysters Special Offer Prompt"* (excerpt from Group 3 flash fiction)

Group 4 worked from the prompt *"Social media lets us unlock lots of different identities...And I find it strange to try and bring these together. I think I am a different person to different people"*. It was written in the form of three diary entries from the perspective of the main protagonist, Crosby, a transgender man

who is now a successful entrepreneur and happy with his professional and personal life. Crosby encounters someone from his past who blackmails him by threatening to make his history of transition known to the public without Crosby's consent by hacking into his old social media accounts.

At the end of the exercise, all groups read their story aloud for the other participants and a concluding debate followed. Interestingly, each group chose a very different register and format to tell their story (fairytale, "thriller", short logs, and diary entries).

All groups keenly engaged in the exercise, and while at the beginning they felt that the time they were given to write the story was quite short, once they began working on the prompts and sharing ideas and inspiration they were able to develop storylines and agree on the plot fairly quickly.

It is important to note that the four stories that were produced elaborated more on issues of *impact* of technology on work and life and on challenges emerging when work and life blend, rather than technological scenarios where technology perhaps behaves in surprising or unusual ways. This could be due to a number of factors: the background and interests of the participants for one, or the earlier part of the workshop, which focused more heavily on the results of the interview study. However, interestingly, the flash fiction exercise seemed to work very well to go deeper into some of the issues that had emerged from the data in terms of people's lives and choices. In other words, they creatively explored and developed some of the overarching themes from the empirical material, although this was not encouraged or prescribed.

Discussion and Conclusions

This was our first time using a fiction technique for this purpose, and, subsequently, using flash fiction. As it was an exploratory exercise, we can only draw some limited insights from the experience, which nonetheless can be useful to us and to others in planning and executing future similar activities.

We found the phrasing of the prompts to be key in shaping the development of the story and its tone. While we chose the prompts on the basis of their potential to generate a surprising story because they included ambiguous words or situations, it is clear that the themes that the quotes brought up were equally resonant with the participants. The fictions that were produced highlighted certain dimensions of the empirical data that we identified through the thematic analysis we had conducted, therefore there was a definite resonance that emerged in the fictions even if the data excerpts were removed from their context and adapted as prompts.

The participants commented on how the exercise helped them to reflect the open questions surrounding the subject area of work/life boundaries. This is not surprising in itself as they were already interested in the topic (hence their participation in the workshop), they did however mention how they developed those themes in ways that they had not imagined before. They particularly engaged with issues of isolation, pressure and frustration, which heavily featured in the stories. Group 4 particularly stressed how the prompt encouraged them to think about deeper and more personal aspects of identity (i.e. gender identity in their story), while up to that point they had been thinking about identity more in terms of online presence (e.g. digital accounts, different approaches to self-presentation online, etc.).

From our perspective as facilitators, the discussion that groups were having while plotting their stories and how the story was developed collaboratively also constituted valuable data. We were able to take notes while observing the groups at work, and to flag important points in their process of plotting and composing the story. Group 2 removed themselves from the workshop room and to another area nearby so that they could play the song "Under Pressure" in the background as they were working to keep them in the mood of the story. The subsequently decided to have the song playing in the background as they read the story aloud as they felt it was an important part of their work.

Overall, the exercise was useful in exploring the potential of flash fiction as a technique to elaborate on data excerpts through fiction and imagination. The rapid response and short frame for the stories pushed the participants to choose which themes to develop, but also prompted their creativity.

Of course, as we mentioned, this was a small and exploratory exercise and it presented several limitations, and therefore it cannot be used as basis for generalisations. The participants were researchers external to our team and were not familiar with the data beforehand. Furthermore, the length of the subsequent discussion was limited due to the workshop constraints, and no follow-up exercise with the same people was possible. Also, the exercise took place in groups (which was also unavoidable due to time constraints), and it would be interesting to see what could emerge from individuals to write their own stories. By choice, we decided to only play the role of facilitators and documenters of the exercise, instead of taking part in the group work. This enabled us to see which themes emerging from the data were elaborated by external people, however it could be very valuable for those researchers who know the data intimately to be engaged in an activity that is outside the more consolidated approaches to analysis.

At the same time, it was interesting to see what people decided to create through the inspiration from data that we knew well. It would be interesting to carefully design and embed a fiction-focused methodology along these lines as part of a research project and with a group fully immersed in the data, to see

which directions the stories would take and which impact it could have on how data is made sense of and interpreted.

The workshop exercise we described in this paper constituted an important moment of reflexivity for us - the researchers. Reflexivity is in our view an essential aspect of conducting qualitative research (Altheide and Johnsen 1994; Gergen and Gergen 2000). In terms of how the exercise impacted our own relationship with the data, a first round of analysis had been completed at the time of the workshop, and (as we mentioned) some of the themes we had identified also emerged in the fictions that the groups created. This surprised us to a certain extent, and also encouraged us to continue our reflection over the data to expand the discussion of such themes. Furthermore, the activity of creating prompts out of the interview transcripts was a valuable exercise as it made us pay attention to some of the nuances in the way participants expressed their view, such as choice of particular words and the humour they put in describing their work/life challenges to us. This encouraged us to look back onto the emotional tone of the answers we received, which is a novel angle for analysis that we hope to develop further in future work.

In conclusion, in this exploratory paper we presented our views on adopting a fiction technique for the collaborative reflection among researchers over qualitative data, and presented what was a very small exercise that only explored using this technique and barely scratched the surface. It is obvious that more substantial exploration of this topic is needed. However, we do believe that reporting our reflections on the experience can spark methodological debate and discussion in the CSCW and HCI communities.

Acknowledgments

This work was supported by the EPSRC Balance Network activity grant "Managing Technologies Around Work and Life". We sincerely thank all the participants in the flash fiction exercise. An early version of this work was presented at the CHI 2017 workshop "Design Fictions for Mixed-Reality Performances"; we thank the organisers and participants for the valuable feedback. Many thanks also to the anonymous ECSCW 2019 reviewers for their helpful comments.

References

Altheide D. L, Johnsen J. M. (1994): 'The reflexive turn in qualitative research', in Denzin N. K, and Lincoln Y. S. (eds.), *Handbook of qualitative research*, Sage, Thousand Oaks, CA, pp. 485–499.

Auger, J. (2013): 'Speculative Design: crafting the speculation', *Digital Creativity*, vol. 24, no. 1, pp. 11-35

Bell, A., Ensslin, A. and Rustad, H. K. (2014) (Eds): *Analyzing Digital Fiction*, Routledge, London

Blythe M., and Wright, P. (2006): 'Pastiche Scenarios: Fiction as a Resource for Experience Centred Design', *Interacting with Computers*, vol. 18, issue 5, September 2006, pp. 1139-1164

Blythe, M. (2014): 'Research through design fiction: narrative in real and imaginary abstracts'. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*, ACM, New York, NY, USA, 703-712. DOI: http://dx.doi.org/10.1145/2556288.2557098

Brown, B., Bleecker, J., D'Adamo, M. et al (2016): 'The IKEA Catalogue: Design Fiction in Academic and Industrial Collaborations', *GROUP'16 Proc. Of the 19th International Conference on Supporting Group Work*, ACM, New York, pp. 335-344

Bødker, S. (2015): 'Third Wave HCI, 10 years later – participation and sharing', *Interactions*, vol. 22, no. 5, pp. 24-31

Cabrero, D. G., Winschiers-Teophilus, H. and Abdelnour-Nocera, J. (2016): 'A Critique of Personas as representations of "the other" in Cross-Cultural Technology Design', in *AfriCHI'16, Proc. Of the First African Conference on Human-Computer Interaction*, Nairobi, Kenya — November 21 - 25, 2016, ACM, New York, 149-154

Carroll, J.M. (1995): *Scenario-Based Design*, John Wiley & Sons, New York

Cheon, E. and Su, N. M. (2017): 'Configuring the User: "Robots have Needs Too"', *CSCW'17 Proc. of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, ACM, New York, pp.191-206

Ciolfi, L. and Lockley, E. (2018): 'From Work to Life and Back Again. Examining the digitally-mediated work/life practices of a group of knowledge workers', *Computer-Supported Cooperative Work*, vol. 33, 2018, no. 5-6, pp. 499-543

Comber, R., Pargman, D., Bates, O. and Eriksson, E. (2018): 'The Futures of Computing and Wisdom', In Proceedings of the 10th Nordic Conference on Human-Computer Interaction - NordiCHI '18, ACM Press, 2018, p. 960-963

DiSalvo, C. (2012): 'Spectacles and Tropes: Speculative Design and Contemporary Food Cultures', *Fibreculture: Special Issue on Networked Utopias and Speculative Futures*, Issue 20

Draper, H. and Sorell, T. (2014): 'Using Robots to Modify Demanding or Impolite Behavior of Older People', in Beetz, M., Johnston, B., Williams, M.A. (eds.), *Social Robotics. ICSR 2014*, Lecture Notes in Computer Science, vol. 8755, Springer, Cham, pp 125-134

Dunne, T. and Raby, F. (2013): *United Micro Kingdoms (UMK) Design Fiction*. The Design Museum, London.

Galef, D. (2016): *Brevity: A Flash Fiction Handbook*. Columbia University Press, New York.

Gaver, W., Dunne, T. and E. Pacenti (1999): 'Design: Cultural Probes', *Interactions*, vol. 6, no. 1,

pp. 21-29

Gergen M. M. and Gergen K. J. (2000): 'Qualitative inquiry: Tensions and transformation', in Denzin N. K, Lincoln Y. S. (eds.). *Handbook of qualitative research. 2nd ed*, Sage, Thousand Oaks, CA, pp. 1025–1046

Grudin, J. (2003): 'The west wing: fiction can serve politics', *Scandinavian Journal of Information Systems*, 15:1, 01/01/2003, pp. 73-77

Hales, D (2013): 'Design Fictions an introduction and provisional taxonomy', *Digital Creativity*, 24:1, pp.1-10, DOI:10.1080/14626268.2013.769453

Iivari, N. and Kuutti, K. (2018): 'Critical design in interaction design and children: impossible, inappropriate or critical imperative?', In *IDC'18 Proceedings of the 17th ACM Conference on Interaction Design and Children*, ACM, New York, pp. 456-464

Kim, Jeong-Hee (2016): *Understanding Narrative Inquiry. The Crafting and Analysis of Stories as Research*, SAGE, Thousand Oaks, CA.

Linehan, C., Kirman, B., Blythe, M., Reeves, S., Wakkary, R., Desjardins, A. and Tenenbaum, J. (2014): 'Alternate endings: using fiction to explore design futures', *Proceedings of the ACM SIGCHI conference on Human Factors in Computing Systems (CHI 2014) extended abstracts*. Toronto: Canada.

Muller, M. and Liao, V. (2017): 'Exploring AI Values and Ethics through Participatory Design Fictions, *HCIC 2017: Design Futures*, available online: http://qveraliao.com/hcic2017.pdf

Putnam, C., Rose, E., Johnson, E.J and Kolko, B (2009): 'Adapting User-Centered Design Methods to Design for Diverse Populations', *Journal of Information Technologies and International Development*, 5: 4 (Winter 2009), pp. 51-73

Richardson, L. (1994): 'Writing: A Method of Inquiry', In N. K. Denzin & Y. S. Lincoln (Eds.): *Handbook of qualitative research*, Sage Publications, Thousand Oaks, CA, pp. 516-529

Setyowati, L. (2016): 'Analysing the students' ability to write opinion essay using flash fiction', *JELTL Journal of English Language Teaching and Linguistics*, 1(1) 2016, pp. 79-91

Shapard, R. (2012): 'The Remarkable Reinvention of Very Short Fiction', *World Literature Today*, 86:5 (Sept/Oct 2012), pp. 46-49

Smit, I., Cila, N. and Lupetti, M. L. (2018): 'Near future cities of things: addressing dilemmas through design fiction', *Proceedings of NordiCHI '18*, ACM, New york, pp. 787-800

Sterling B. (2009): 'Design Fictions', *Interactions*, vol 16. issue 3, pp.

Wright, P.C. and McCarthy, J.C. (2004): 'The value of the novel in designing for experience', in A. Pirhonen, C. Roast, P. Saariluoma, H, Isom (eds.): *Future Interaction Design*. Springer, Amsterdam, pp. 9-30

Medina Angarita, M. A.; Nolte, A. (2019): Does it matter why we hack? – Exploring the impact of goal alignment in hackathons. In: Proceedings of the 17th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing and the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep01

Does it matter why we hack? – Exploring the impact of goal alignment in hackathons

Maria Angelica Medina Angarita[1], Alexander Nolte[1,2]
University of Tartu[1], Carnegie Mellon University[2]
{*maria.medina, alexander.nolte*}*@ut.ee*

Abstract. Time-bounded events such as hackathons have become increasingly popular in recent years. During these events participants typically form teams, exercise fast prototype development, challenge themselves to innovate, practice new skills, collaborate with diverse team members, and compete against other teams. Hackathon organizers have a certain vision in mind about which outcome they would like to achieve and design the event based on this vision. Participants on the other hand do not necessarily share the same vision and come with their own goals and aspirations. While work in related fields suggests that it is important for goals of organizers and participants to align in order to achieve them this might be different in hackathons due to their unique setup. Drawing from literature we identified potential goals of organizers and participants and conducted a case study of three hackathons focusing on the alignment of goals between organizers and participants. Our findings indicate that the goals of organizers and participants did not align in all cases, that goal awareness on the part of the organizers appears may have a stronger impact on goal achievement and that hackathons appear to have inherent characteristics that can materialize even when not planned for.

Introduction

Hackathons are time-bounded events during which participants gather in teams and attempt to complete a project of interest to them (Pe-Than et al., 2019). Originating from coding competitions in the early 2000s, such events have garnered increased interest from both practitioners and researchers as evident by the large number of

global events taking place every weekend[1] and the emergence of academic events focusing on the topic (Pe-Than et al., 2018). This increase in interest has led hackathons to proliferated into various domains ranging from corporations conducting internal hackathons (Nolte et al., 2018) and higher education institutions (Kienzler and Fontanesi, 2017) to civic engagement groups (Hartmann et al., 2018; Henderson, 2015) and (online) communities (Angelidis et al., 2016). Within those domains, individuals organize hackathons with different goals in mind such as public engagement to raise awareness and advocacy (Taylor and Clarke, 2018), tackling civic and environmental issues (Baccarne et al., 2014; Porter et al., 2017), fostering innovation (Briscoe and Mulligan, 2014), creating technology (Stoltzfus et al., 2017), expanding or creating networks of interested individuals (Möller et al., 2014), spreading knowledge about technologies (Nandi and Mandernach, 2016) and others.

The aforementioned goals are often communicated to potentially interested individuals prior to the hackathon in the form of marketing material which contains a short summary of the overall theme of the hackathon as well as core organizational details such as place and time. More detailed information is typically delivered in the form of an introductory presentation at the event including *"an overview of the event, any rules and regulations, and themes and goals"* (Decker et al., 2015).

The reasons for participants to go to a hackathon, however, do not necessarily match those of the organizers. While participants might share similar goals such as learning, inducing social change, building a product and finding a team to work with, they sometimes also participate in hackathons for glory, free pizza, finding employment and winning prices (Briscoe and Mulligan, 2014). In the context of game jams for example, fun is a key reasons for (re-)attendance (Arya et al., 2013). This points towards a potential disparity between the goals of organizers and participants that has not been investigated in depth in existing work on hackathons. We address this gap by asking the following research question:

RQ1: How do the goals of hackathon organizers and participants align?

Moreover, it is not clear whether it is inevitably necessary for goals of organizers and participants to be aligned in order for both groups to achieve them and to perceive a hackathon as a satisfying experience. There are hints towards the necessity of goals alignment in the work conducted by Hou and Wang (2017) in the context of a civic data hackathon. They found tension between two intertwined goals: helping with data driven work and learning with the purpose of getting involved in the work of NPOs. Conflicts in this case were resolved by brokers. Literature on project management also suggests that goal alignment is important

[1] Hackathons organized by the largest hackathon league alone register more than 65.000 students among more than 200 events each year (MLH, https://mlh.io/about)

for project success (Skulmoski and Hartman, 1999) and that goal alignment enables the achievement of performance outcomes (Stephen and Coote, 2007). Similarly, work in the context of work groups suggests that *"a basic coordination problem in the management of groups is to increase alignment of personal goals with the group goals"* (Zhang and Chiu, 2012), pointing out that it is important for individuals to share their goals with their group and achieve goal alignment to succeed. Correspondingly, goal misalignment has been found to cause conflict within groups in the context of joint software reviews where issue resolution can be affected by differences between goals of different reviewers (Kingston et al., 2000). These contexts however are considerably different from hackathons in that work groups members are bound by contracts and common social norms while this is not the case in hackathons where participants are not necessarily familiar with each other before the event. In addition, hackathons might have inherent characteristics that might foster certain goals simply due to the format, such as networking as pointed out by Drouhard et al. (2016). To further investigate this aspect, we will also ask the following research question:

RQ2: How does goal alignment influence goal achievement at hackathons?

In order to answer these two research questions, we conducted a qualitative case study covering three hackathons. Our results indicate that organizers and participants of the hackathons we studied shared some common goals such as networking and learning. Digging deeper, however, we found that the specifics of these goals to be different between organizers and participants e.g. related to being interested in learning different skills. We also found indication that goal alignment was not necessarily a prerequisite for goal achievement, but instead, goal awareness could improve goal achievement. We also found indications for the hackathon format having inherent characteristics which can contribute to the achievement of certain goals without explicit planning.

Hackathon goals

There are a number of reasons why individuals organize and participate in hackathons as pointed out in the introduction. Based on a review of relevant literature in IEEE Explorer, ACM Digital Library and Semantic Scholar, we developed a coding scheme that covers goals for hackathons in various contexts (c.f. Table 1 for an overview). These goals can be roughly divided into professional (A) and personal goals (B). We consider goals as professional when they can directly influence the career of an individual such as learning a specific skill this individual can use during her/his everyday work. In addition to the goals we identified from related work we discovered additional goals during our analysis. We will discuss them together in the following.

One goal commonly found in hackathons is networking (Briscoe and Mulligan, 2014) which can be broken down into professional networking (A3) with the aim to further an individual's career (A5) or into a personal goal to meet new people (B1). Learning is also often cited as a motivation for individuals to organize and participate in hackathons (Saravi et al., 2018) since hackathons have been found to support knowledge exchange (Ghouila et al., 2018) and foster collaborative learning (Porras et al., 2018). Learning can again be perceived as a professional (A4) or personal goal (B5).

Hackathons are also often organized in the context of entrepreneurship (Beltrán, 2017). It thus common for participants of hackathons to focus on creating a prototype (A1) and founding a start-up after a hackathon has ended (A2). Furthermore, it might be interesting for them to see what other participants are working on (A6). Moreover, individuals with a specific start-up idea in mind might also want to seek potential investors (A7) or individuals that are interested in working together with them (A8). All of the aforementioned goals are related to the professional development of the respective participants.

Hackathons are however not only a means of promoting individual careers and developing start-up companies. Participants also often come to a hackathon because they are fun (B4) events (Calco and Veeck, 2015), because participants are interested in the experience (B3), or they perceive it to be a personal challenge (B2).

Table 1. Coding scheme

A Professional Goals	Source
A1 Developing an idea into a prototype	Briscoe and Mulligan, 2014; Trainer et al., 2016
A2 Creating a startup	Cobham et al., 2017; Decker et al., 2015
A3 Networking	Briscoe and Mulligan, 2014; Nandi and Mandernach, 2016
A4 Learning	Briscoe and Mulligan, 2014; Ghouila et al., 2018
A5 Professional development	Cobham et al., 2017
A6 Seeing new ideas	Deducted from analysis
A7 Investment	Briscoe and Mulligan, 2014
A8 HR	Briscoe and Mulligan, 2014
B Personal Goals	
B1 Meeting new people	Komssi et al., 2015; Taylor and Clarke, 2018
B2 Personal challenge	Deducted from analysis
B3 Having a new experience	Deducted from analysis

B4 Having fun	Arya et al., 2013; Calco and Veeck, 2015; Saravi et al., 2018
B5 Learning	Nandi and Mandernach, 2016; Porras et al., 2018

Study setting

To answer the research questions described in the introduction we conducted a case study of three different hackathons in two Northern European countries (c.f. Figure 1 for some impressions). We selected hackathons that were similar in scope in terms of number of days, number of participants and type of audience (c.f. table 2 for an overview). The type of hackathon we studied was catalytic (Drouhard et al., 2016). The style of the work environment was competitive, and teams could win prizes that would allow them to continue working on their projects after the hackathon had ended. However, didactic talks, professional development and the pursuit of impact were also part of the hackathons.

Table 2. Hackathon anatomy

Hackathon	H1	H2	H3
Duration	48 hours	48 hours	48 hours
Number of Participants	~40	37	36
Participants	Researchers, students, entrepreneurs	Students, entrepreneurs	Students, enthusiasts

The theme of hackathon 1 (H1) was to develop innovative bio-technical products with the possibility of winning prizes that would allow teams to continue working on their projects after the hackathon. This weekend long event was attended by more than 40 students, researchers and entrepreneurs. It started with design workshop held by the organizers before the participants began working on their ideas and prototypes. Hackathon 2 (H2) focused on sustainability and ecological impact. This weekend long event hosted 37 participants including students and entrepreneurs who developed prototypes and competed for prizes that would allow them to continue working on their projects. Hackathon 3 (H3) was part of a larger effort in that similar hackathons with the same theme organized by the same group of people took place simultaneously in over 100 locations. H3 aimed to solve data visualization, hardware and other prototyping challenges related to space exploration. During this weekend long hackathon, 36 participants including students and enthusiasts gathered in teams and collaborated with each other to develop technical solutions for the aforementioned challenges. Each hackathon thus had the development of a technical artifact at its core.

Figure 1. Stills of hackathon 1 (top right), hackathon 2 (left), and hackathon 3 (bottom right).

Research methods

We conducted semi-structured retrospective interviews with organizers and participants at each aforementioned hackathons. This approach appeared to be feasible since we are interested in studying the perception of participants and organizers of hackathons on their individual goals and whether or not they have been achieved. Similar designs have been successfully applied in other exploratory studies on hackathons (Page et al., 2016; Nolte et al., 2018).

For the interviews we developed an interview script focusing on goal alignment and goal achievement. The themes of the interview were:

- Goals: The aims of hackathon organizers and participants related to their careers and their personal interests (e.g. *What were your professional goals for this hackathon?*)
- Goal assessment: The metrics participants and organizers applied to assess their goal achievement (e.g. *What goals did you achieve?*)
- Technology: The tools participants used to cooperate with each other. (e.g. *What tools did you use to collaborate with your teammates?*)
- Hackathon attendance: How many times participants have been to a hackathon before (e.g. *Is it your first time at a hackathon?*)
- Hackathon sustainability: Whether participants are planning to continue working in their projects after the event has ended (e.g. *Do you think you will continue working on your idea?*)
- Background information: Educational and professional history (e.g. *Tell me about your educational background.*)

The interview script was piloted with one hackathon participant and one organizer. Based on this pretest we adjusted the interview script to ensure the

feasibility, flow and appropriateness of the questions. We selected at least one organizers and multiple participants for our study. The selection of suitable participants was based on their background (students, entrepreneurs), hackathon experience (first timers and experienced hackathon participants), locality (individuals that live in a place for a long time and individuals that recently moved) and whether or not they pitched an idea at the hackathon (c.f. Table 3 for an overview).

Table 3. Demographic profile of the participants and organizers

Hackathon	Students	Entrepreneurs	First timers	Locals	Idea pitched	Organizers
H1	P1,P2		P1			O1
H2	P3	P1, P5, P6, P8	P2, P3, P7	P2, P4	P6	O1, O2
H3	P1, P2, P4		P1	P3, P4		O1

After transcribing all interviews one of the authors manually coded the interviews using the coding scheme we derived from literature (c.f. Table 1). We followed a deductive coding procedure starting with the pre-defined codes adding categories if necessary (e.g. *Personal challenge* (B2) in Table 1).

Goal alignment and achievement of hackathon organizers and participants

During the course of this section we will first elaborate on the goals of hackathon organizers (O) and participants (P) of each hackathon (H1, H2, H3) based on our coding scheme (c.f. Table 1). We will then elaborate on their alignment within one hackathon (RQ1) and the potential impact of the alignment on whether or not goals were achieved (RQ2). Overall, we found that organizers and participants did not interact with each other on a regular basis. The organizers mainly focused on the operation of the hackathon making sure that e.g. the planned schedule would be followed. Interaction between organizers and participants during the event was limited to participants asking individual organizers specific questions e.g. about upcoming activities. Organizers mostly reached participants for coordination purposes during the event personally (H1, H3) and used Slack (H2).

Teams internally mainly communicated in person using other tools such as GoogleDrive or Facebook messages mainly to share files. Each team could decide on their own toolset with no interference by the organizers.

Goal alignment and achievement between the organizer and participants of hackathon 1 (H1)

The main aim for the organizer (O1) was "*to provide the platform for the people that work in this area, for them to get together*" (O1), by creating an environment for participants to work on their ideas. S/he also aimed for the participants to expand their network (A3) and to acquire new skills (A4).

The participants, in comparison, mentioned networking (A3) and learning (A4) as their main goals. For example, P1 mentioned that s/he wanted to "*meet people, speak to them, understand what their point of view is on problems*" (P1). It thus appears as if participants and organizer goals were aligned since both aimed for participants to expand their own networks and acquire new skills. However, when looking deeper into those two aspects we identified a number of differences.

For the organizer (O1), networking (A3) meant "*to connect students, just beginners, or early stage student teams with the local startup network"* (O1), and to boost the generation and implementation of ideas related to the theme of the hackathon. O1 particularly aimed to connect participants with specialists working at an entrepreneurial center where they could find advice and tools to continue working on their projects. For participants, however, networking was not linked to identifying individuals that would support them in continuing to work on their project. For them, networking was rather related to learning. P1 mentioned for example that s/he "*just wanted to learn new things from new people*" (P1).

Similar to networking (A3), we found that learning (A4) initially appeared to be a mutually shared goal for organizers and participants. However, when looking deeper we also found that the organizer and participants aimed for different learning aspects. For the organizer it was important that the participating researchers would learn how to pitch because s/he thought that "*researchers tend to be too complicated*" (O1). The organizer also aimed for the participants to learn about design thinking (*"this whole empathy creating with the potential user or customer"*, O1).

Conversely, participants wanted to learn about the theme of the hackathon. P1 wanted to learn for her/his professional development "*there is some innovation in biology which I am searching for, and I really want to take part in it"* (P1); and P2 was interested in "*how we can, for example, improve our lives to be better and to live longer*" (P2).

From the previous analysis it becomes clear that there is a disparity between organizer and participant goals related to networking (A3) and learning (A4). This appeared to mainly affect goal achievement on the part of the organizer, since participants reported to have achieved their goals, for example, P2 was able to learn about patients with Parkinson's disease, "*for me, it was like a discovery that we can actually help these people*" (P2).

Goal alignment and achievement between the organizers and participants of hackathon 2 (H2)

For H2, the main goal for the organizers (O1, O2) was "*to connect [country1] and [country2] people who work in tech or in the topic, with the end goal of having more businesses run by both [country1] and [country2]*" (O1). In general, the organizers thus aimed for the participants to network (A3) by meeting new people (B1) and then form teams to develop an idea into a prototype (A1), which could potentially lead to creating a new startup (A2). To foster this last goal, they *"invite[d], like, angel investors, so yeah, we give them the tools and it's always up to the participants to like use those tools*" (O2) thus supporting them to find investors (A7).

The goals of the participants however were much more diverse. Most participants mentioned that they were interested in learning (A4), P5, P6 and P8 mentioned wanting to develop an idea into a prototype (A1), P1 and P8 were eager about seeing new ideas (A6). P1 and P5 aimed to find investing opportunities and investment (A7), P1 hoped to find potential future employees i.e. achieve HR (A8), P7 and P8 were looking for a new experience (B3) and P1, P6 and P8 participated for fun (B4). The aim of participants related to learning was generally to learn "*something new*" (P5) by working with teammates (P4) or by talking to people at the hackathon (P6). One participant also wanted to learn more about how to create a start-up (P2) and improve her/his presentation skills (P2).

All participants reported that they achieved their respective with a few notable exceptions: P2 reported that s/he did not manage to learn what s/he aimed to learn, P1 nor P5 did not find investment opportunities, not investors, (A7), and P1 was not able to achieve HR (A8) by finding potential employees. Finding investors and investing opportunities – a mutually shared goal between organizers and participants – was thus not achieved.

Most participants mentioned that they were partially able to achieve their learning goals while pointing towards multiple potential reasons for not achieving them. One participant mentioned that *"it's […] very difficult to learn a new skill in two days*" (P7) while another participant stated that *"there's always room to learn more"* (P4). Next to these general remarks P2 also stated that it was not possible for her/him to improve her/his presentation skills because someone else in her/his team was in charge of pitching. In addition, s/he stated that s/he would have expected to be taught more about e.g. how to write a business plan to create a start-up (P2). It would have certainly been possible for the organizers to support these participants to achieve their goals by planning the hackathon in a different way. There was thus no direct misalignment between participant and organizer goals but rather a lack of awareness about specific participant goals on the part of the organizers which might have resulted in some participants not being able to achieve their learning goals.

Another issue we found was that one participant was not able to work on her/his idea because s/he did not find a team and *"s/he didn't feel so great about any of the other ideas so […] s/he just decided to leave"* (P6). This could have also been something that could have been spotted by the hackathon organizers especially since one of their goals was to support participants to turn their idea into a prototype.

For other goals of the organizers such as teams actually creating a start-up it is not possible to assess them at the end of the hackathon since they need to be assessed long term.

Goal alignment and achievement between the organizer and participants of hackathon 3 (H3)

For the organizer of H3 the main goal was to create an environment for people where they could network (A3) and collaborate on their project ideas (A1). The participants mentioned that their goals included networking (A3), learning (A4), meeting new people (B1), having a personal challenge (B2), experiencing something new (B3), and having fun (B4). Both organizers and participants thus mentioned networking as one of their primary goals. However, compared to both previous hackathons, there was not disparity in the respective details of this goal. Both participants and organizers aimed to foster professional networking with the aim to support the professional ambitions of the participants.

The organizer mentioned that her/his goal related to networking (A3) might only have partially been achieved. This perception was based on her/him expecting students to get together in their free time ("*if you think that only the students between each other will do projects, activities together, then that would be nice*", O1). Participants however were excited about meeting new peers and potentially starting long term relationships, for example, P4 wanted to "*see more people in my field, make connections*" (P4), and P2 commented that "*maybe some other time I need advice*" and s/he could get it from the people s/he met at the hackathon (P2).

Apart from meeting new people, participants were also eager about learning, having a new experience and a personal challenge. P1 and P2 reported they were able to achieve these goals, meanwhile, P3 and P4 reported to have achieved all of them, except for learning (A4). P3 mentioned that s/he wanted to learn more about public speaking but also noted that her/his anxiety *"won't go away in one second*" (P3) but rather would *"get better, like, day by day,*" (P3). Finally, P4 wanted to learn about the hiring processes in companies but eventually did not ask the mentors – who were recruited by the organizers from local companies – about it. This is certainly something that the hackathon organizers could foster if they would be aware of it.

Discussion

The previously described analysis reveals a number of interesting aspects related to the question how the goals of hackathon organizers and participants align (RQ1). Our findings first indicate that the goals of participants and organizers mainly align with respect to networking and learning. Other goals such as fostering the creation of start-up companies (A2) were more important for organizers while finding investments (A7) and having fun (B4) were more important for participants.
However, when looking closer we found that participants and organizers were often interested in different aspects of networking and learning despite them both frequently mentioning these two goals. Organizers mainly focused on professional networking (A3) while participants were mainly interested in getting to know people on a personal level (B1). Similarly, when it comes to learning, participants on one hand were interested in learning about a large variety of different aspects such as creating a start-up, pitching, learning about new ideas and learning about how to collaborate with a group of people. Organizers on the other hand mainly focused on pitching, and although they were present throughout the entire duration of the hackathons, they mainly focused on facilitating operations and making sure "*that everything went smoothly*" (O2, H2). They only interacted with participants when triggered by them. The goals of organizers and participants thus appear to be well aligned at first sight but were not particularly well aligned when breaking them down into different aspects of e.g. learning.

Despite this apparent lack of alignment between the goals of organizers and participants we did however find that most participants reported to have achieved their goals (RQ2). The goals that they achieved were mainly related to aspects such as having fun (B4), learning about something new or improving existing skills both professionally (A4) and personally (B5). The specific aspects of learning that they reported to have achieved however differed not only between different participants but also between participants and organizers.
Our analysis also revealed that participants in the same team did not necessarily share the same goals. Moreover, each team created their own communication and coordination strategy including the decision which technologies they would use to communicate and exchange artifacts during the hackathon. These findings are similar to the ones reported by Trainer et al. (2016) and Lundbjerg et al. (2017). It should also be noted that teams rarely used technology to communicate. They did however use tools such as Google Drive and Slack to share artifacts. Ensuring awareness about tasks and goals was thus fostered by the co-located setting rather than additional technologies.

The fact that most participants reported to have achieved their respective goals despite an apparent lack of alignment points to the assumption that some goals are simply inherent to the nature of hackathons which means that it might not be required to specifically plan for them. Learning and networking are the two main

examples for this. Both can – according to our study – be achieved simply due to the nature of hackathons in that people that do not necessarily know each other before coming together during a hackathon to work on a project idea. Such ideas often involve working on something that is not necessarily familiar to all team members which in turn requires individuals to acquire new skills in order to complete their project and to pitch their project idea to an audience. This finding is in line with previous work by Warner and Guo (2017) who found that learning for participants can be incidental (as a consequence of doing), opportunistic (by taking advantage of the tools and facilities), or from talking to peers and that learning can thus be an inherent hackathon characteristic. Similar findings were reported by Drouhard et al. (2016).

That being said we also identified situations in which participants did not achieve goals such as attracting investment (A7) and creating a start-up (A2) directly. These specific goals however are very unlikely to be achieved during a hackathon and should thus be assessed in the months after the event has ended. In such cases organizers could point out that such goals are unrealistic and that a hackathon can be a starting point on a longer journey but that reaching these goals requires longer term investment. This is in line with previous work by Komssi et al. (2015) who stated that *"hackathons by themselves don't initiate new business, they require mechanisms in place in order to commercialize their results"*.

We also found situations in which participants did not achieve their specific learning goals despite them having the possibility to do so. One participant wanted to learn about pitching but someone else in the team pitched their idea instead, one participant wanted to learn more about start-up creation but there was no specific advice during the hackathon. Another participant wanted to learn about the hiring process in companies but did not get to talk to hackathon mentors about it. Those goals could probably have been achieved if the organizers would have been aware of them and adjusted the procedure during the hackathon. This points towards **goal awareness** on the part of the organizers being more important than actual goal alignment. To foster goal awareness organizers could in the future e.g. approach participants and ask them about what they would like to achieve during the hackathon. This would also organizers to support participants reaching their goals. The hackathon format itself however provides an opportunity for social interaction that inherently fosters goals such as networking and learning.

Contrary to Hou and Wang (2017) we did not find any tensions being created by misaligned goals. Our findings thus also stand in contrast to work in the context of project management where goal alignment is considered to be an important prerequisite for project success (Skulmoski and Hartman, 1999) and misalignment can lead to conflict (Kingston et al., 2000). This contrast however might stem from the fact that in our case participants in particular were mainly focused on learning and networking rather than completing a particular project. Both of these goals can

be achieved by individuals during a hackathon without any specific external support as discussed before.

Limitations

The aim of this study was to identify goals of hackathon organizers and participants, their alignment and the potential effects of goal alignment on their achievement. This particular phenomenon has received limited attention in research so far. It thus appeared reasonable to conduct an in-depth case study. We do however acknowledge that despite developing and applying a coding scheme that is grounded in relevant literature and carefully selecting study participants studying groups is different hackathons working on different problems with different goals might yield different results.

Future work

Based on the results of this study our aim is to develop a framework of goals which will serve as a basis for a survey instrument to study the interdependence of the different identified goals on a larger scale. Our sample for this study will include similar participants to those we studied thus covering individuals who are going to hackathons for the first time, individuals who have been to many hackathons, individuals who have ideas that they want to work on during a hackathon and individuals who do not. For the upcoming study we will also adjust our research focus by including the aspect of goal awareness as discussed in the previous section. We will also use the identified goals as a basis for a of keywords to conduct a quantitative case study on a larger hackathon database. These two studies combined will allow us to identify how different goals can influence hackathon outcomes as well as the perception of hackathon outcomes by participants.

Acknowledgments

Thanks to all interviewees who donated their time for the sake of this research.

References

Angelidis, P., Berman, L., Casas-Perez, M. de la L., Celi, L. A., Dafoulas, G. E., Dagan, A., … Winkler, E. (2016): 'The hackathon model to spur innovation around global mHealth'. *Journal of Medical Engineering & Technology*, vol. 40(7-8), pp. 392-399.

Arya, A., Chastine, J., Preston, J., & Fowler, A. (2013): 'An International Study on Learning and Process Choices in the Global Game Jam'. *International Journal of Game-Based Learning*, vol. 3, pp. 27-46.

Baccarne, B., Mechant, P., Schuurman, D. Colpaert, P. & De Marez, L. (2014): 'Urban socio-technical innovations with and by citizens'. *Interdisciplinary Studies Journal*, vol. 3, no. 4, pp. 143-156.

Beltrán, H. (2017): 'Staging the Hackathon: Codeworlds and Code Work in México'. ISSI GRADUATE FELLOWS WORKING PAPER SERIES 2016-2017.78. UC Berkeley: Institute for the Study of Societal Issues.

Briscoe, G. and Mulligan, C. (2014): 'Digital Innovation: The Hackathon Phenomenon'. *Creativeworks London*, vol. 6, pp. 1–13.

Calco, M. and Veeck, A. (2015): 'The Markathon: Adapting the Hackathon Model for an Introductory Marketing Class Project'. *Marketing Education Review*, vol. 25, no. 1, pp. 33-38.

Cobham, D., Gowen, C., Hargrave, B.J., Jacques, K.W., Laurel J. & Ringham, S. (2017): 'From hackathon to student enterprise: an evaluation of creating successful and sustainable student entrepreneurial activity initiated by a university hackathon'. *In 9th annual International Conference on Education and New Learning Technologies. EDULEARN.*

Decker, A., Eiselt, K., and Voll, K. (2015): 'Understanding and improving the culture of hackathons: Think global hack local'. *IEEE Frontiers in Education Conference (FIE)*, pp. 1-8.

Drouhard, M., Tanweer, A. and Fiore-Gartland, B. (2016): 'A typology of hackathon events'. In *Hacking at Time-Bound Events Workshop at Computer Supported Cooperative Work*, pp. 1-4.

Ghouila, A., Siwo, G. H., Entfellner, J.-B. D., Panji, S., Button-Simons, K. A., Davis, S. Z., … Mulder, N. (2018): 'Hackathons as a means of accelerating scientific discoveries and knowledge transfer'. In *Genome Research,* vol. 28, no.5, pp. 759–765

Hartmann, S., Mainka, A. and Stock, W. G. (2018): 'Innovation Contests: How to Engage Citizens in Solving Urban Problems? Enhancing Knowledge Discovery and Innovation in the Digital Era'. *IGI Global*, pp. 254–273.

Henderson, S. (2015): 'Getting the most out of hackathons for social good'. In *Volunteer Engagement 2.0: Ideas and insights changing the world,* pp. 182–194.

Hou, Y. and Wang, D. (2017): 'Hacking with NPOs: Collaborative Analytics and Broker Roles in Civic Data Hackathons'. *Proceedings of the ACM on Human-Computer Interaction,* vol. 1.

Kienzler, H. and Fontanesi, C. (2017): 'Learning through inquiry: A global health hackathon'. *Teaching in Higher Education*, vol. 22, no. 2, pp. 129–142.

Kingston, G., Jeffery, R. and Huang, W. (2000): 'An explanatory study on the goal alignment problem in joint software reviews'. In *Proceedings 2000 Australian Software Engineering Conference*, pp. 63-72.

Komssi, M., Pichlis, D., Raatikainen, M., Kindström, K. and Järvinen, J. (2015): 'What are Hackathons for?'. In *IEEE Software*, vol. 32, no. 5, pp. 60-67.

Lundbjerg, E.H., Osten, J.P., Kanto, R. & Bjørn, P. (2017): 'The Hackerspace Manifested as a DIY-IoT Entity: Shaping and Protecting the Identity of the Community'. ECSCW Exploratory Papers.

Möller, S., Afgan, E., Banck, M., Bonnal, R. J., Booth, T., Chilton, J., … Chapman, B. A. (2014): 'Community-driven development for computational biology at Sprints, Hackathons and Codefests'. *BMC Bioinformatics*, 15(S14), S7.

Nandi, A. and Mandernach, M. (2016): 'Hackathons as an informal learning platform'. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education (SIGCSE '16),* pp. 346–351.

Nolte, A., Pe-Than, E. P. P., Filippova, A., Bird, C., Scallen, S., & Herbsleb, J. D. (2018): 'You Hacked and Now What? Exploring Outcomes of a Corporate Hackathon'. *Proceedings of the ACM on Human-Computer Interaction*, vol. 2, no. CSCW, pp. 129:1–129:23.

Page, F., Sweeney, S., Bruce, F. and Baxter, S. (2016): 'The use of hackathon in design education: an opportunistic exploration'. *Proceedings of the 18th International Conference on Engineering and Product Design Education (E&PDE)*, pp. 246-251.

Pe-Than, E.P.P., Herbsleb, J., Nolte, A., Gerber, E., Fiore-Gartland, B., Chapman, B., Moser, A., and Wilkins-Diehr, N. (2018): 'The 2nd Workshop on Hacking and Making at Time-Bounded Events: Current Trends and Next Steps in Research and Event Design'. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems.* ACM., New York, NY, USA, Paper W35, 8 pages.

Pe-Than, E.P.P., Nolte, A., Filippova, A., Bird, C., Scallen, S. and Herbsleb, J.D. (2019): 'Designing Corporate Hackathons With a Purpose'. *IEEE Software*, vol. 36, no. 1, pp. 15–22.

Porras, J., Khakurel, J., Ikonen, J., Happonen, A., Knutas, A., Herala, A., and Drögehorn, O. (2018): 'Hackathons in software engineering education: lessons learned from a decade of events'. In *Proceedings of the 2nd International Workshop on Software Engineering Education for Millennials (SEEM '18),* pp. 40-47.

Porter, E., Bopp, C., Gerber E. and Voida, A. (2017): 'Reappropriating Hackathons: The Production Work of the CHI4Good Day of Service'. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. pp. 810–814, ACM.

Saravi, S., Joannou, D., Kalawsky, R. S., King, M. R. N., Marr, I., Hall, M., … Hill, A. (2018): 'A Systems Engineering Hackathon – A Methodology Involving Multiple Stakeholders to Progress Conceptual Design of a Complex Engineered Product', *IEEE Access*, vol. 6, pp. 38399-38410.

Skulmoski, G. J., and Hartman, F.T. (1999): 'Project alignment: The key to successful cost engineering', *AACE International*. Transactions of the Annual Meeting, pp. PM41-PM45.

Stephen, A. T., and Coote, L.V. (2007): 'Interfirm behavior and goal alignment in relational exchanges', *Journal of Business Research*, Elsevier, vol. 60, no. 4, pp. 285-295,

Stoltzfus, A., Rosenberg, M., Lapp, H., Budd, A., Cranston, K., Pontelli, E., … Vos, R. A. (2017): 'Community and Code: Nine Lessons from Nine NESCent Hackathons', *F1000Research*, vol. 6, pp. 786.

Taylor, N., and Clarke, L. (2018): 'Everybody's Hacking: Participation and the Mainstreaming of Hackathons', *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 172, ACM.

Trainer, E.H., Kalyanasundaram, A., Chaihirunkarn, C. & Herbsleb, J.D. (2016): 'How to Hackathon: Socio-technical Tradeoffs in Brief, Intensive Collocation.', CSCW.

Warner, J., and Guo, P. J. (2017): 'Hack.edu: Examining How College Hackathons Are Perceived By Student Attendees and Non-Attendees'. In *Proceedings of the 2017 ACM Conference on International Computing Education Research (ICER '17),* pp. 254-262.

Zhang, Y., and Chiu, C. (2012): 'Goal commitment and alignment of personal goals predict group identification only when the goals are shared', *Group Processes & Intergroup Relations*, vol. 15, no. 3, pp. 425–437.

Prabhat, S.; Motwani, A.; Rangaswamy, N. (2019): Longitudinal analysis of a #boycott movement on Indian online platforms: Case of collective action and online boycott. In: Proceedings of the 17th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing an the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep10

Longitudinal analysis of a #boycott movement on Indian online platforms: Case of collective action and online boycott

Shantanu Prabhat, Aditya Motwani, Nimmi Rangaswamy
International Institute of Information Technology - Hyderabad, India
shantanu.prabhat@research.iiit.ac.in,aditya.motwani@research.iiit.ac.in, nimmi.rangaswamy@iiit.ac.in

Abstract.

Lately, Indian online platforms have witnessed recurring consumer boycott calls in forms of uninstalling/down-voting applications on the app stores. For our exploratory study, we conduct a longitudinal analysis of one of these controversies involving online boycotting of an e-commerce company (Snapdeal) over a controversial statement issued by their brand ambassador (Aamir Khan) which hurt religious and nationalist sentiments of users on Twitter. Through the lens of this study, we try to understand emergent collective user behaviour and how collective action has begun to play out in online (troll) communities. We call the user behaviour in our study as trolling behaviour as - the call to boycott appears in order to silence an influential voice which challenges a nationalist narrative of these users. A broader implication of such behaviour seems to be strong arming any counter narrative with a threat of potential backlash and financial harm. This analysis is important as online deviant and trolling behaviour by group of users is increasingly influencing socio-political agendas online. It contributes to broader CSCW understanding of online platforms and collective behaviour. We also situate our work in online consumer boycotting behaviour.

Introduction

Deviance and trolling behaviour is increasingly becoming common on online platforms (Sanfilippo et.al, 2017, Flores et.al, 2018). Members identifying with certain beliefs (both bots and actual users), try to employ their strength in numbers to drive political agendas and frame narratives online. Users engaging in this behaviour collectively mobilise and often co-ordinate acts such as rigging online polls, targeting and harassing specific individuals, altering dialogue in different online communities (Shachaf et.al, 2010, Massanari et.al). The infamous #Gamergate on Reddit, stands as an example of misogynist activism that perpetuated harmful anti-feminist narratives by reddit users. A case of collective harassment played out on the platform, targeting women (and other minority) game designers, developers and journalists. Other such studied instances in recent times have been infiltration of Russian trolls (on both sides of political spectrum) during the US Presidential election to influence discourse (Badawy et.al, 2016), mass raiding of subreddit communities (Kumar et.al, 2018), and collective down-voting of reviews motivated by political causes [1]. All these instances underpin the act of collective activity to make a political point online.

We observe that these users engaging in such acts could be either loosely. or strongly connected to each other on online platforms (Bennett et.al, 2012). They could be part of dedicated subreddit communities, or loosely connected through ways of “hashtags" and topics of interest. But they seem to be overcoming problems that underpin executing successful collective and co-ordinated action (Flores et.al, 2018). These problems often happen to be - lack of effective messaging, non-participation and free-riders in the community, lack of mobilisation or challenge in informing members in the network about developing agendas and subsequent strategies.(McClain et.al, 2017, Obregon et.al, 2017, Piven et.al, 1991).

Our work seeks to understand a similar and recurring form of online collective action that is playing out on Indian online platforms. Users mobilise over Twitter to uninstall/down-vote an app endorsed by a certain celebrity they have ideological disagreements with. Broadly, it is a call to boycott a product the ‘controversial’ celebrity is endorsing and leverage power as consumers to build pressure on the corporate entity to disassociate from the celebrity, lest, they face economic consequences.
While product boycotts are not uncommon (Hawkins,2010, Klein,2004, Li, 2018) as a political message, social media is increasingly being utilised for online call for boycotts such as the Netflix boycott, #deleteUber [2], #boycottAmazon. In the Indian context, we identified four such (major) events where a celebrity made a

[1] https://www.hollywoodreporter.com/heat-vision/captain-marvel-rotten-tomatoes-audience-score-sandbagged-by-trolls-1193280

[2] https://www.nytimes.com/2017/01/31/business/delete-uber.html

controversial statement which a section of the twitter populace took objection to and in-turn made a call to boycott the product endorsed by the celebrity. While these calls were general product boycott calls, they also involved uninstalling and/or down-voting the app on Google and Apple's app-stores. These events targeted Snapdeal (an e-commerce platform) , Amazon (e-commerce) [3], Republic Tv (news channel app) [4], and Snapchat (social networking app) . Of these events, we choose the Snapdeal boycott event for purpose of our exploratory analysis. The Snapdeal boycott was the first major online boycott that played out in the Indian online space with a long event cycle of outrage and boycott. This gives us opportunities to understand emergent collective user behaviour and how collective action has begun to play out in online (troll) communities.

While consumer boycotts are usually initiated over ethical reasons such as boycotting companies for underpaying their staff (Klein et.al, 2004) , or boycotting oligopolies as it happened in Morocco , we call the online boycotts in our study as largely trolling behaviour. This is because - through acts of down voting and boycotting, they are trying to undermine and silence the voice of anyone who challenges their socio-political narrative. It suits a larger agenda of creating a standard narrative online, and putting celebrity figures in a fear of potential backlash and financial harm. The online accounts involved in such behaviour rarely engaged in discourse and discussion, and were largely targeting (trolling) a certain celebrity.

To this end, we investigate the following research questions-
RQ1)- What are the behavioral patterns of anti-Aamir Khan/Snapdeal users that are indicative of collective action?
RQ2)- Which tweets resonated amongst these users for mobilisation and for collective action?

With understanding from these, we position our findings on how collective behaviour emerges and is executed among users engaged in trolling behaviour. Our work is also situated in understanding online boycott actions and what how it can inform similar studies.

Background

We situate our work in consumer boycott behaviour, use of social media for collective action, and contemporary research on challenges in collective action.

[3] https://www.indiatimes.com/entertainment/people-delete-amazon-app-for-its-association-with-swara-bhaskar-boycottamazon-trends-343882.html

[4] https://www.inuth.com/india/assuming-arnab-goswami-called-keralites-shameless-netizens-sink-channels-rating

Consumer boycott behaviour

Boycotts have long been used a tool to build pressure, and leverage power as consumers to force companies to make different choices in their business policies. Though the efficacy of this tactic is contested (Sen et.al, 2001, Neilson et.al), but that has not stopped different activist groups from issuing calls to boycott over issues of environment, fair pay, ethical treatment of animals amongst other popular issues (Klein et.al, 2004). The website, www.ethicalconsumer.org keeps a track of companies which have been called to boycott over different "ethical" issues. Another such phenomena of buycotts involves actively choosing to buy products of the rival company that one is boycotting. It has been studied that for a lay person, building a list of brands that one can choose from seems to be more effective than building a list of brands that one can't choose from (Neilson et.al).

Seeing this as a CSCW problem, recent research has focused on overcoming some of the challenges that underline organising successful boycotts. Li et.al (2018) designed a light and semi-automated prototype *out of site* which eases boycotting goods by restricting web searches through browser extensions. Mills et.al (2015) analysed the effective ways in which users on Reddit could oppose and boycott SOPA's (Stop Online Privacy Act) provisions and build an information base. We situate our work in a similar CSCW lens of understanding calls to boycott through network and content analysis.

Social media and collective action

Social media has been instrumental in recent movements such as the Occupy Wall Street, Black Lives Matter, Arab Spring in mobilisation efforts. Live tweeting and reporting through social media were considered more reliable than newspaper reports (Grossman, 2018). Faster dissemination of information through social media, also helped in better mobilisation and co-ordination of activities (McClain et.al, 2017). Through social media, people with similar stake and activist goals could come together to pool their resources and devise strategies(Schradie, 2018). While not everyone was an active participant and some were labeled as slacktivists too (Lee et.al, 2013) (only online participation and no on ground contribution), social media helped reach out to a wider array of people with quick, and very personal messaging.

Online platforms are being utilised to not just mobilise for offline collective actions, but people are engaging online with events that have larger socio-political implications. Online petition signing has grown as a medium to signal support with hopes of an on-ground effect (Hale et.al, 2013). There has been similar research on designing systems for collective activities for civic tasks (Cheng et.al, 2014) .The flip side of this happen to be instances of deviant acts on these online platforms such as collective doxxing and harassing individuals and groups(Massanari, 2015), collectively altering reviews of books(Bhaskar, 2015) and apps online, and fudging and altering results of online polls. Study by Flores et.al(2018)on r/The_Donald subreddit identified behaviour patterns of the most active participants and calls to

action which influence and engaged the participants of the community the most. Massanari.(2015), studied how reddit users collectively came together to doxx and harass female journalists and gamers. Thus, social media hasn't been a value neutral platform and is being engaged with for a variety of civic and political causes. These incidents highlight the importance of studying how narratives are built and sustained on online platforms for setting larger political agendas.

Challenges in mobilisation and collective action

Research in sociology and allied domains have highlighted the problems underpinning successful mobilisation leading to collective action. Collective action theory(McClain et.al, 2017) highlights the free rider problem- wherein the benefits of collective action are shared by all but participants are unsure on who will put in the effort for the same. There is a constant conflict between personal and collective goals which play out in such contexts. Other studies(Obregon et.al, 2017, Piven et.al, 1991, Choudhary et.al, 2016) highlight how effective communication must take center stage where all members of the group are aware of the next course of action to take, and the benefits that they gather from it. Participants are also more likely to contribute to a cause if they are aware that the movement is more likely to succeed. (McClain et.al, 2017). We discuss our findings and what it informs us on the ability of the users participating in #boycott movement to overcome some of these challenges for collective action.

Timeline

We provide a brief timeline on the events as they unfolded in the Snapdeal-Aamir Khan controversy.

23 Nov 2015: Aamir Khan, a popular film celebrity from India, makes a remark on the "growing intolerance in India" and how his wife (Kiran Rao) suggested moving out.

24 Nov 2015: Statement picked up by the media and Twitter, leads to outrage over the said remark. "Nationalist" sentiments claim to be hurt over the remark.

25 Nov 2015: Twitter users troll Aamir Khan and begin a call to boycott Snapdeal (a brand endorsed by Aamir Khan). Users begin to write poor reviews, giving poor rating and uninstalling the app with the hope to leverage power as consumers and build pressure over Snapdeal to remove Aamir Khan.

7 Jan 2016: Aamir Khan loses Govt. of India's Incredible India ambassador contract

5 Feb 2016: Snapdeal does not renew Aamir Khan's contract.

Data and methods

We gather our data for our longitudinal analysis by running the Twitter API and employing other search and retrieval techniques such as using Twitter Scraper [5] - an open source front-end retrieval tool which performs queries on Twitter's advanced search platform. Since, the twitter API doesn't retrieve all the data, it was important to additionally pivot to other methods of extraction.

We gather tweets for the controversy from 20 November 2015 to 3 March 2016, using a seeding process. We initiate gathering tweets with seed hashtags of #AamirKhan and #Intolerance. As and when the tweets come, we increase our set of hashtags. We also ran boolean queries such as "aamir AND snapdeal", "aamir OR intolerance". It is also noted that running a search in the scraper or the API returned results for substrings and were non case sensitive. Searching for "#Aamir" returns results for aamir, AAMIR, #Aamir.

In order to ensure that the tweets used in our analysis reflect our lists in our topics of interest, we ran it through our own post-filtering process. We converted the tweet into a lower case string, and tokenized them using Stanford's NLTK library (Manning). We run our tweets through a regular expression built to check if the tweet contained at least one of our keywords and hashtags. All these tweets were then selected in our data set.

Data set description

The data set consisted of 117632 number of tweets, by 63452 number of users. Of these 52127 number of tweets were NOT retweets (but include quote retweeting). This doesn't imply that tweets which were not retweets were unique tweets, as we also observe that there were tweets which were copied across handles (indicating co-ordination and spamming).

RQ1: Patterns indicating Collective Behaviour

To analyse the behaviour that are indicative of collective and coordinated actions, we explore the affordance provided by Twitter to users and how they got appropriated in a collective setting. In our exploratory study, we analyse the @-mentioning behaviour (users tagging other users) and hashtags used and operated at scale. We also analyse evidence for possible marginalisation and amplification of specific voices on the platform.

@-mentioning behaviour

To study this, we plot the in-degrees and out-degrees of all user handles on a scatter plot as shown in fig-I. In-degree for an account are the number of user handles

[5] https://github.com/taspinar/twitterscraper

mentioning them, and out-degree refers to the number of user handles that *they mention*.

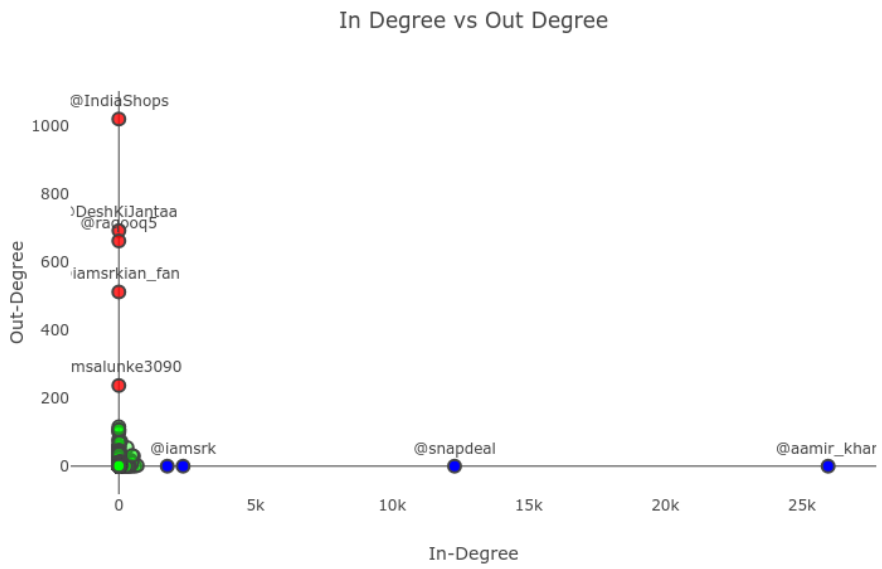


Figure 1. Indegree - Outdegree for user handles.

From fig-I we see that both Snapdeal and Aamir Khan have very high in-degrees(26134 and 12674, respectively) and 0 out-degree. Hence, of all the users who were in the fold of controversy - neither Aamir or Snapdeal replied to them but they were targeted very heavily. The other group of outliers we see are the handles with 0 in-degree but very high out-degrees. We identified 10 such handles. On careful examination of the user profiles, 4 of them appeared to be bots(DFR, 2017) (guidelines - only retweets, same tweet multiple times). They were massively tagging @snapdeal, and @aamirkhan and news media handles such as @ndtv. The other 6 accounts, didn't appear to be bots but were consistently tagging @snapdeal, @aamirkan and other a couple of other users by drawing their attention to the controversy. On an average each such handle was tagging 50-60 other user handles. The third kind of user handles that we see in the figure are the ones with an average indegree of 2 (σ = 0.2) and an out-degree of 5 (σ = 10). These were about 97% of the total user handles. We observe that there is very effective targeting that the users are engaging by ways of @-mentioning. Twitter mentions are usually conversational in nature. By mentioning @-userhandles, users draw attention of one another to a particular tweet. Here we find that instead of an intended conversation, the "@" + username mentions were used to identify the intended targets and employ what we call as *collective targeting*. There is a clear establishment of who the targets are, and this information is being disseminated in the network.

Hashtags as messages and communities

We pick the top 13 hashtags which account for about 98% of the tweets and analyse
1) What is the lexical nature of the hashtags and how are they employed?
2) How overlapping are the communities which employ a certain kind of hashtag and what does it tell us?

For these 13 hashtags, two authors of these papers inductively coded the nature of the hashtags. We outline the guidelines that we followed on coding the hashtags in the different categories. Categories of the hashtags, and their percentages are displayed in table-I.

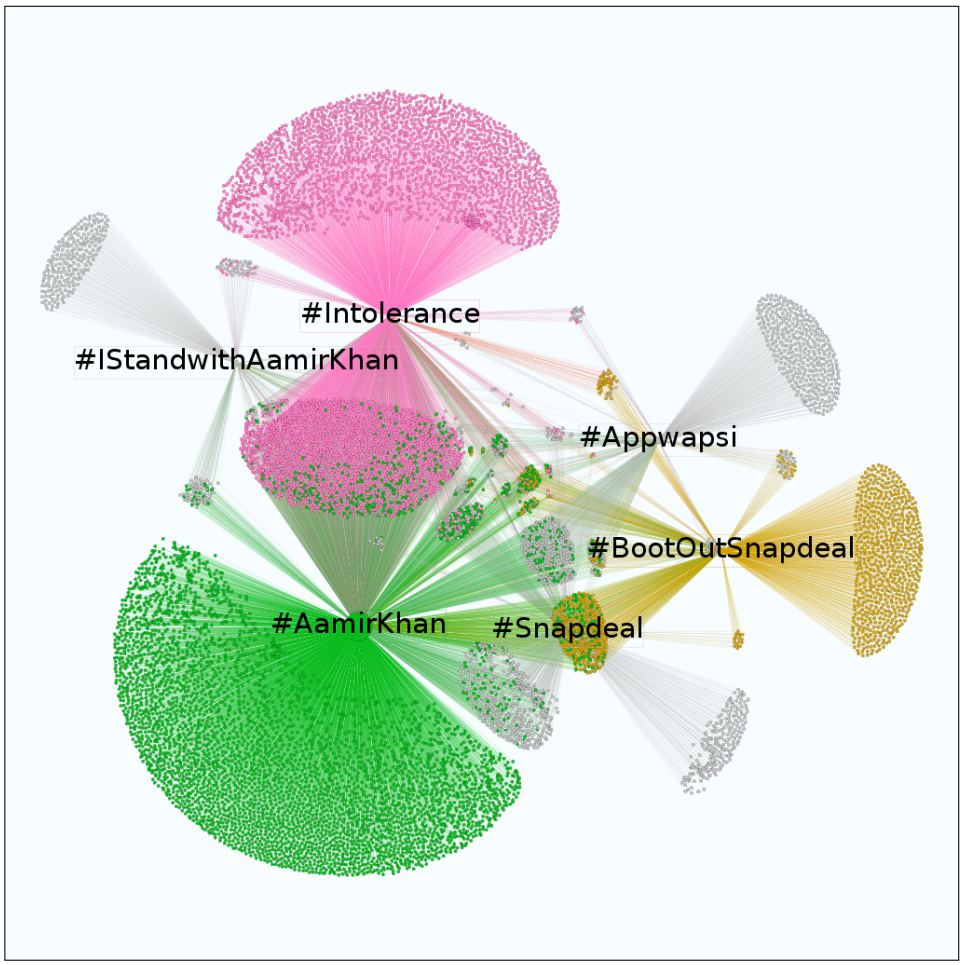


Figure 2. Ego networks for 6 hashtags. 2 from each category.

Calls for action hashtags: Hashtags which fell into this category contained verbs directing the people to do an act of rebel such as "bootout", "appwaapsi (return the app)", "(Say)No", "uninstall", "boycott". About 42% of tweets employed the use of such a hashtag.

Category of hashtag	Hashtags	Percentage
Call for action	Notosnapdeal, boycottSnapdeal, bootoutSnapdeal, AppWaapsi, shameaamir	42%
Opinionated	IstandWithAamir, AamirRightOrWrong, GetWellSoonPK AamirInsultsIndia	21%
Topical	Intolerance, AamirKhan, Snapdeal, Intolerancedebate	82.1%

Table I. Hashtag categories and hashtags.

Opinionated hashtags: The lexical nature of these hashtags was such that they expressed an opinion on the issue.For instance hashtag such as #AamirInsultsIndia is an opinionated hashtag. Only about 21% of user used one of these hashtags.

Topical hashtags These hashtags were neutral and expressed a large topic of interest in the controversy such as #Aamir, #Snapdeal. About 82.1% of tweets contained at least one hashtag from this category.

We also build ego networks for all the top 20 hashtags. At the center of each ego network is the hashtag, and the nodes are all the user handles who used the particular hashtag. For easy representation purposes, we pick the top 2 hashtags from each category.

We observe that #AamirKhan and #Intolerance are both large ego networks and are also overlapping. Thus, the sentiment of *intolerance* was made to ride over *Aamir Khan* and build a strong association between the two entities. Hashtags aren't used to merely indicate topics of discussion but also to disseminate information on the platform. We also observe that there is a significant(63.2%) overlap between #BootoutSnapdeal (call for action hashtag) and other hashtags. Similarly, other calls for action hashtags such as #Appwapsi, #BootoutSnapdeal all co-occur with other topical hashtags such as (#Snapdeal, #Intolerance) and have overlap of greater than 50%.

Use of hashtags creates small and temporary communities between the users of the said hashtag, with each other. In such a mobilising setting, we see that hashtags are employed to build and inform association of one entity to the other such as - Aamir Khan and intolerance. Hashtag use with significant overlap with use of another hashtag indicate *piggybacking*, where one bit of information seems to ride of another closely associated information and is informed around the social network. Calls for action, which are commonly observed in collective action settings, are informed here through means of *call for action* hashtags.

Marginalising and Othering

We analyse whether there were user behaviour patterns which were indicative of marginalising the alternate voices, and increasing tweet volume content to amplify the voice of the group calling for boycotts.

We filter the tweets' contents which were more than 50% similar to each other. These tweets were possibly made by bots, or by users running multiple accounts

looking to increase the volume count of tweets in the controversy. We find that there are 12 such spammed messages spanning 38 user accounts. There is also higher than average URL usage amongst the tweets (compared by running twitter stream API for an hour and comparing it to the 1% of tweets captured).

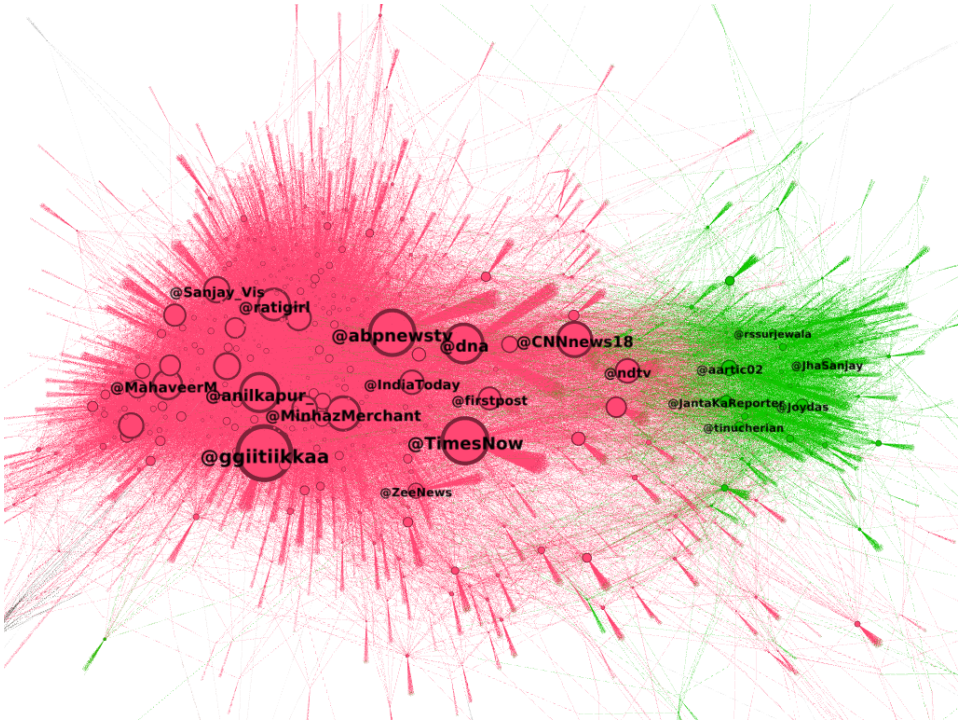


Figure 3. Bipartition and polarity in network through retweet patterns..

Since retweets are methods of amplification, and also method of dissemination of information through the network to users with like minded agenda (such as followers) (Zaman, 2010), retweeting patterns give us insights into which group of users cluster together and how "loud" is their voice. It is also indicative of polarity on the platform (Garimella et.al, 2018). Polarity also leads to marginalisation of alternate voices. We analyse if there was polarity on the twitter platform during the controversy and what kind of voices were amplified.

We use a graph partitioning algorithm METIS(Garimella et.al, 2018, Karypis et.al, 1995) which gives us partitions in the network based on a controversy score. We feed to it the retweet pattern between 2 users. If user1 retweeeted user2, there was a direct edge between them. The algorithm partitions the network into two largest connected components with 71.63% nodes and 26.22% nodes. We visualise an undirected graph using a force directed algorithm Gephi's ForceAtlas2 [6].

From the figure III we observe that the largest connected component on the left is much bigger (71.63% nodes v/s 26.22%) and is much denser too. For purpose of our exploratory study, we pick the top 0.02% handles which make up approximately

[6] https://github.com/gephi/gephi/wiki/Force-Atlas-2

73% of the total tweet volume in the network. These come out to be about 65 handles. Analysing these handles we find that, 7 handles were of news-outlets. 49 of these handles were engaged in anti-Snapdeal, anti-Aamir Khan sentiment and were a part of the cluster on the left. Only 9 of the most retweeted handles spoke either in favour of Aamir Khan, Snapdeal or were neutral in their stance. The

The bigger, denser cluster on left is the one engaged in collective boycotting behaviour. We observe that they are more interconnected, and amplifying each others' voices. This marginalises and shunts the alternate voice, and possibly sets the agenda on the platform.

RQ2: Analysing effective mobilising and calls for action messaging

To study the kind of arguments and calls for action, that resonated the most with people who possibly took part in some level of action of reviewing/uninstalling/commenting, we identify such user accounts. From our list of user handles from our twitter dataset we identify users who also wrote reviews and gave poor rating on the appstore. We scraped this list with the belief that if a user handle is @manoj_kumar, and they go by the name of Manoj_Kumar on Appstore reviews, they are likely to be same as they engaged in similar activity of *tweeting* in Aamir-Snapdeal controversy as well as *uninstalling/downvoting* as a part of the controversy. We use Heedzy [7] to gather data for app reviews.

We also identified user handles (by means of regular expression search string) which made claims of uninstalling on twitter such as *"Just uninstalled Snapdeal"*. and added it to our list of users who possibly engaged in collective action behaviour. A lot of such users also attached a screenshot along with their tweet. We do not claim that this list is exhaustive or without errors, but it is useful as an exploratory analysis.

Analysing the arguments based tweets which resonated with the users:

Methods - *argument based tweets*

Since this is a controversy where different arguments are posed to challenge and direct narratives (against or pro Aamir Khan and Snapdeal), we aim to analyse what broad arguments appeared in our dataset. To do this aim, we conduct a thematic analysis of tweets and code them deductively using a "logical fallacy" framework popular in STS studies (Copi, 1953). We call these tweets-based frames and then calculate user engagement metrics through RTs and Favoriting of such tweets. The two authors of this paper conducted the deductive analysis inspired by the framework. We narrowed the arguments that appear in our dataset into, what came up as six logical fallacies - (1) Denial (2) False Dilemma (3) False

[7] https://heedzy.com/

Equivalence (4) Suspicion (5) Whataboutery and (6) Anecdotal. They were agreed upon with a cohen's kappa of 0.79. The number of tweets relevant to our study were 23981 tweets. We coded only a subsample of these tweets which received at least one RT or 'favourite' and were 3216 in number.

We outline each category and present a representative tweet alongside.

Denial: Tweets in this category made an argument that if they (the user) didn't feel that intolerance existed, then Aamir's claim of intolerance was wrong.

> "There is no intolerance as I can't see it."

False Dilemma: These tweets presented a false comparison in which one must choose.

> "If there is intolerance then he should move to another country (Pakistan)"

False Equivalence: Such tweets made an unjust equivalence between *being Hindu* (his wife) and *being safe* because India is a Hindu majority country.

> "Aamir Khan who has a Hindu wife, can't feel unsafe in a Hindu majority country."

Suspicion: These tweets cast suspicion over motives of Aamir Khan's statement. Usual suspicions were over movie promotions, publicity or being agent of the opposition party.

> "The intolerance remark is for publicity."

Whataboutery: Such tweets engaged in whataboutery, a popular political tactic. These tweets aimed to shift the discourse from Muslim minority being under threat to cases when Hindu majority are unsafe.

> "What about intolerance when Hindus are attacked."

Anecdotal: These tweets sought to undermine the intolerance statement by citing an anecdotal evidence.

> "If he (Aamir Khan) can freely cite his opinion, then there can't be intolerance."

Category of Argument	Engagement
Denial	**Retweets** μ: 156 | σ: 32.05 **Favourites** μ: 321.07 | σ: 17
False Dilemma	**Retweets** μ: 340 | σ: 46.08 **Favourites** μ: 266.78 | σ: 15.6
False Equivalence	**Retweets** μ: 18.02 | σ: 2.31 **Favourites** μ: 24.8 | σ: 3.98
Suspicion	**Retweets** μ: 9 | σ: 3.6 **Favourites** μ: 42.29 | σ: 7.8
Whataboutery	**Retweets** μ: 91 | σ: 2.3 **Favourites** μ: 187 | σ: 21.3
Anecdotal	**Retweets** μ: 228 | σ: 34.8 **Favourites** μ: 119 | σ: 23.1

Table II. Argument category and engagement stats.

Results - *arguments based tweets*

As we are interested in the critical in the arguments which set the tone for mobilisation and resonated the most with the boycotters, we analyse engagement metrics in terms of retweets and favourites. The descriptive statistics for all the arguments and their engagements are in table - II.

We see that *denial* and *anecdotal* arguments resonated the most in terms of both retweets and favourite counts. There was a statistically significant difference between the groups for retweets (ANOVA (F = 21.02, p=0.029)) and favourite counts (ANOVA (F=41.2, p=0.041)).

This could be explained by their deep held beliefs which deny the existence of intolerance in the country, and their anecdotal experiences of not facing any consequences of effects of religious animosity. These arguments hence align with their nationalist world view. This probably also encourages them to partake in action of boycotting to challenge the growing narrative of intolerance (propagated by Aamir Khan) by uninstalling/down-voting.

Amongst retweets we see that, tweets in *False dilemma* were highly retweeted. Tweets under false dilemma where Aamir Khan was expected to make a choice between moving out of the country or accepting things the way they are, had very strong in their language too. This could explain the high retweet behaviour, as retweeting as a practice is also linked to the emotion conveyed in the tweet (Svelch et.al, 2016)

Amongst the tweets which were favourited, we observe that tweets under *Whataboutery* were the ones which were highly favourited. *Whataboutery* as a political tactic is very common and resonates with people as a defensive response. This high favoriting behaviour could be explained by this. However, tweets which were framed as a Suspicion over Aamir Khan's motives of making the statement weren't received very well(neither in terms of RTs or favourites). It could be believed that people on Twitter aren't doubting the intention of making the statement but engaging with the argument itself and countering it with frames such as of denial of his experience, and countering with their own anecdotal experiences.

Analysing calls to action tweets:

Method - *Calls to action tweets*

We borrow the literature from Fleishman(1988) and thematically and deductively classify our call to action tweets into three categories of Direct call strategy, Progress visibility strategy, Solidarity strategy.
Cohen's kappa was 0.59 with an inter-rater reliability of 71.3%. The tweets under study were 42337 tweets. We code only a sub sample of these tweets which received at least one RT or one 'favourite' and were 4317 in number.

Direct Strategy: These tweets made direct calls for action, and clearly outlined the steps on how to make an effective boycott statement by uninstalling/down-voting the Snapdeal app on the playstore.

> "Go to Play Store, Select @snapdeal and rate them 1 Star * and comment that it is only because of @aamirkhan."

Progress visibility strategy: Tweets in this category were the ones which were indicating real(or, fake) progress on the uninstalling and down-voting process. These tweets were intended to present a picture that a successful boycott movement was happening, and others must take part too.

> "Wow 85,000 people angry with #AamirKhan's hypocrisy, uninstalled @snapdeal app. #AppWapsi will hurt badly!"

Solidarity strategy: Tweets in this category seemed to indicate a solidarity amongst people who were against Aamir Khan, Snapdeal and had pro-nationalist sentiments and what they were supposed to do to avenge it.

"People who are united against #AAMIRKHAN statement must uninstall @Snapdeal."

Results- *on calls to action*

We calculate which calls to action resonated the most with this set of users (the ones who engaged in some level of boycott and uninstalling behaviour). We employ the use of retweets and their favourites to understand their engagement.

In terms of retweets, we find that most of the users engaged with *direct strategy* at least once. 49.02% of users retweeted at least one tweet from this category. In terms of favouriting, we see that the engagement is much lower than retweets for this category. Only 26% favourited at least one tweet from the *direct strategy*. This is possibly because, while retweeting a direct call might reach new users and encourage them to partake in the boycott, favouriting doesn't achieve the same goal.

Around 38.09% of users retweeted at least one tweet announcing some real (or unreal) progress of the movement(*progress visibility strategy*) while 37.2% favourited it. We see that the *progress visibility strategy*, trumps over the *direct strategy* in favouriting.

Only 13.9% of these users retweeted tweets indicating solidarity (*solidarity strategy*). The favouriting count however was the highest amongst all other categories at 41.2%. It suggests that while users might not see much merit in retweeting a tweet with a call to action in terms of solidarity, the sentiment resonates with them.

Category of Argument	% users having RT at least once	% users having Favorited at least once	Median number of Users
Direct Strategy	49.02 %	26 %	**Retweets** Median: 521 **Favourites** Median: 121
Progress Visibility Strategy	38.09 %	37.2 %	**Retweets** Median: 91 **Favourites** Median: 140
Solidarity Strategy	13.9 %	41.2 %	**Retweets** Median: 340 **Favourites** Median: 221

Table III. Descriptive statistics for different frames in calls to action.

Table outlines the descriptive statistics for each call to action styles. The median number of users who retweeted the tweets in each category stand at - *Direct*- 521, *Solidarity*-340 , *Progress indicating* - 91 .We find that there was a significant difference between the strategies ANOVA [$F = 33.09$, $p<0.002$]. Direct strategy seems to have the highest engagement across all strategies.

In favouriting behaviour, the median number of users who retweeted the tweets in each category stand at - *Direct*- 521, *Solidarity*-340 , *Progress indicating* - 91 .We find that there was a significant difference between the strategies ANOVA [$F = 45.02$, $p<0.002$]. Solidarity appears to be the most engaging amongst all the strategies.

Discussion

Through our work we cast a lens on how users in online platforms are able to mobilise, troll and utilise their strength in numbers for collective action goals. Learning from our findings inform us of how affordances provided by twitter were used to overcome challenges and set the agenda for collective action, and what kind of messaging best resonated with the active boycotters.

Appropriation of twitter

The like minded users and trolls seem to have appropriated the platform to disseminate information about their goals and co-ordinate their activities. There is a clear targeting of Snapdeal and Aamir Khan, by way of @-mentioning. They are being @- mentioned to draw their attention to the issue and take a stand, and also to inform the rest of the people in the fold of the controversy to build pressure on them by this tactic. People aren't debating the merits or demerits of the controversy with each other, but are isolating the two entities for further targeting through @-mentions. It acts as a way to very clearly establish who the opponents are. Information around mobilisation, is also being disseminated through the use of hashtag. Hashtags are both emotive (#AamirinsultsIndia), and also calls to action (#boycottSnapdeal). They also often co-occur with hashtags used by a large number of people such as #Aamir and #Intolerance which are general hashtags. This tactic seems to be effective in promoting new and growing hashtags such as ones calling for boycotts, and bringing them into the center stage of the controversy. They also serve as mobilising grounds as they bring people using different but similar hashtags together - forming mini topical communities on the platform. Similarly, such controversies also polarise the platform as has been well documented by studies (Zaman, 2010). This polarisation was utilised to amplify voice of boycotts, targeting and shunting down of alternate voices. This is effective in setting agendas and works as a useful pressure building tactic.

Effective messaging

By identifying the users who actively partook in some degree of boycotting behaviour, we analyse messaging strategies that resonate with active boycotters for mobilisation and collective action. Prior work of Flores et.al(2018), which studied the troll community of r/The_Donald, analysed engagement of different calls to action but there was no evidence on whether the users who engaged with it were actually the ones who took part in some degree of collective action behaviour. Similarly, another work of Savage et.al(2018) which discusses the different calls to action strategies deployed by bots on twitter, doesn't measure engagement by users actually boycotting or taking part in some activism behaviour. In our study by identifying such users we are able to gather data on what constitutes effective messaging. As we also see in Savage et.al(2008), direct calls to action were the

most effective and were engaged the most with active boycotters. We also observe that tweets indicating (real/unreal) progress in the movement were effective as such messaging incentives people to participate and mitigates the risk of a 'wasted effort'. In terms of mobilisation, as also observed previously by Svelch et.al(2016), emotive messages were highly favourited and thus, must have resonated highly. But it were the tweets challenging the narrative of intolerant India and rejecting the version peddled by Aamir, were the ones which were the most retweeted by the active boycotters. It can be suggested that the desire to challenge this narrative must have found its way through down-voting/uninstalling.

Replications in #boycott movements

In the Indian context, Snapdeal-Aamir Khan was the first instance that played out in these terms- target the corporate attachment of the celebrity and leverage consumer power through social media in terms of App reviews/uninstalls. Since then it has set a precedence for other similar events to follow. Other events which have played out in Indian context such as the Amazon-Swara Bhaskar [8], (#boycottAmazon), Republic TV - Arnab Goswami [9] (#boycottRepublic), Snapchat- Evan Spiegel [10] (#boycottSnapchat) have employed similar tactics of building pressure and leverage.

However, the event cycles for these events happen to be shorter, possibly explained by a learned behaviour of the people on the platform on how to effectively make a political statement through boycotting the apps and giving poor ratings.

As Li et.al(2018) mention, the only successful online boycott in the context of United States happened to be the #deleteUber movement. This informs us that making a statement through app uninstalls and ratings make a bigger statement than buycotts(buying goods of the rival company that one is boycotting), or generally denouncing company's products without any way to measure on the degree of economic harm being caused. A poor rating on an app is easily visible to everyone (to users and to the company), and it is likely that a corporate entity notices and takes steps to remedy the damaged reputation. Uninstalling and downvoting also happen to be relatively lower effort situations than *actually* boycotting the product or mobilising on the ground. This also explains the success of such tactics.

[8] https://www.firstpost.com/entertainment/boycottamazon-becomes-top-trend-on-twitter-company-down-voted-for-associating-with-swara-bhasker-on-campaign-4441577.html

[9] https://gulfnews.com/world/asia/india/indian-journalist-arnab-goswami-trolled-for-calling-keralites-shameless-1.2270743

[10] https://www.dailyo.in/variety/deleting-snapdeal-boycott-snapchat-ceo-poor-india-digital-racism/story/1/16713.html

Conclusion and limitation

Through an exploratory study of an online boycott movement initiated by the religious and nationalist users on Indian twitter, we shed light on how collective, coordinated action is executed by - appropriating affordances provided by twitter as well as effective messaging to reach out and engage people towards act of collective action. Our study has implications for the broader CSCW community on understanding how messaging for effective collective action works, and how collective trolling behaviour and platform appropriation occurs. As online platforms are becoming place for civic and political discussions, it is crucial to understand when and how platforms and users turn to deviant behaviour.

Our study is limited and can be improved by building a more causal analysis between action and events that follow and finding more micro patterns in collective behaviour. Retweet network on the platform and it's growth traced over time can inform us the way polarisation and amplification occur over time. While our study covers one event traced over a long time for analysis, we can find broader patterns and trends from analysing multiple such occurrences and online boycotts.

References

Badawy, A., E. Ferrara, and K. Lerman (2018): 'Analyzing the Digital Traces of Political Manipulation: The 2016 Russian Interference Twitter Campaign'. In: *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. pp. 258–265.

Bennett, W. L. and A. Segerberg (2012): 'THE LOGIC OF CONNECTIVE ACTION'. *Information, Communication & Society*, vol. 15, no. 5, pp. 739–768.

Cheng, J. and M. Bernstein (2014): 'Catalyst: Triggering Collective Action with Thresholds'. In: *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*. New York, NY, USA, pp. 1211–1221, ACM.

Cheng, J., C. Danescu-Niculescu-Mizil, and J. Leskovec (2015): 'Antisocial Behavior in Online Discussion Communities'. *CoRR*, vol. abs/1504.00680.

Choudhury, M. D., S. Jhaver, B. Sugar, and I. Weber (2016): 'Social Media Participation in an Activist Movement for Racial Equality'. In: *Proceedings of the Tenth International Conference on Web and Social Media, Cologne, Germany, May 17-20, 2016.* pp. 92–101.

Copi, I. M. (1953): *Introduction to Logic*. Pearson/Prentice Hall.

Erickson, M. (2007): 'Review: Andrew Chadwick, Internet Politics: States, Citizens, and New Communication Technologies. Oxford and New York: Oxford University Press, 2006. Âč24.99. 400 pp'. *European Journal of Communication*, vol. 22, no. 2, pp. 237–239.

Fleishman, J. A. (1988): 'The Effects of Decision Framing and Others' Behavior on Cooperation in a Social Dilemma'. *Journal of Conflict Resolution*, vol. 32, no. 1, pp. 162–180.

Flores-Saviaga, C., B. C. Keegan, and S. Savage. 'Mobilizing the Trump Train: Understanding Collective Action in a Political Trolling Community'.

Garimella, K., G. De Francisci Morales, A. Gionis, and M. Mathioudakis (2018a): 'Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship'. In: *Proceedings of the 2018 World Wide Web Conference*. Republic and Canton of Geneva, Switzerland, pp. 913–922, International World Wide Web Conferences Steering Committee.

Garimella, K., G. D. F. Morales, A. Gionis, and M. Mathioudakis (2018b): 'Quantifying controversy on social media'. *ACM Transactions on Social Computing*, vol. 1, no. 1, pp. 3.

Hale, S. A., H. Margetts, and T. Yasseri (2013): 'Petition Growth and Success Rates on the UK No. 10 Downing Street Website'. In: *Proceedings of the 5th Annual ACM Web Science Conference*. New York, NY, USA, pp. 132–138, ACM.

Hawkins, R. A. (2010): 'Boycotts, buycotts and consumer activism in a global context: An overview'. *Management & Organizational History*, vol. 5, no. 2, pp. 123–143.

Karypis, G. and V. Kumar (1995): 'METIS – Unstructured Graph Partitioning and Sparse Matrix Ordering System, Version 2.0'. Technical report.

Klein, J. G., N. C. Smith, and A. John (2004): 'Why We Boycott: Consumer Motivations for Boycott Participation'. *Journal of Marketing*, vol. 68, no. 3, pp. 92–109.

Kumar, S., W. L. Hamilton, J. Leskovec, and D. Jurafsky (2018): 'Community Interaction and Conflict on the Web'. In: *Proceedings of the 2018 World Wide Web Conference*. Republic and Canton of Geneva, Switzerland, pp. 933–943, International World Wide Web Conferences Steering Committee.

lab, D. 'BotSpot: Twelve Ways to Spot a Bot'.

Lee, Y.-H. and G. Hsieh (2013): 'Does Slacktivism Hurt Activism?: The Effects of Moral Balancing and Consistency in Online Activism'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 811–820, ACM.

Li, H., B. Alarcon, S. Milkes Espinosa, and B. Hecht (2018): 'Out of Site: Empowering a New Approach to Online Boycotts'. *Proc. ACM Hum.-Comput. Interact.*, vol. 2, no. CSCW, pp. 106:1–106:28.

Manning, C. D., M. Surdeanu, J. Bauer, J. Finkel, S. J. Bethard, and D. McClosky (2014): 'The Stanford CoreNLP Natural Language Processing Toolkit'. In: *Association for Computational Linguistics (ACL) System Demonstrations*. pp. 55–60.

Massanari, A. (2017): '#Gamergate and The Fappening: How RedditâĂŹs algorithm, governance, and culture support toxic technocultures'. *New Media & Society*, vol. 19, no. 3, pp. 329–346.

McClain Brown, C. (2017): 'Tweets and Mobilisation: Collective Action Theory and Social Media'. *Medijske studije*, vol. 8, pp. 3–22.

Mills, R. and A. Fish (2015): 'A Computational Study of How and Why reddit.com was an Effective Platform in the Campaign Against SOPA'. In: G. Meiselwitz (ed.): *Social Computing and Social Media*. Cham, pp. 229–241, Springer International Publishing.

Neilson, L. A. (2010): 'Boycott or buycott? Understanding political consumerism'. *Journal of Consumer Behaviour*, vol. 9, no. 3, pp. 214–227.

Obregon, R. and T. Tufte (2017): 'Communication, Social Movements, and Collective Action: Toward a New Research Agenda in Communication for Development and Social Change'. *Journal of Communication*, vol. 67, no. 5, pp. 635–645.

Piven, F. F. and R. A. Cloward (1991): 'Collective Protest: A Critique of Resource Mobilization Theory'. *International Journal of Politics, Culture, and Society*, vol. 4, no. 4, pp. 435–458.

Sanfilippo, M., S. Yang, and P. Fichman (2017): 'Managing Online Trolling: From Deviant to Social and Political Trolls'. In: *HICSS*.

Savage, S., A. Monroy-Hernández, and T. Höllerer (2015): 'Botivist: Calling Volunteers to Action Using Online Bots'. *CoRR*, vol. abs/1509.06026.

Sen, S., V. Morwitz, and Z. GÃijrhan-Canli (2001): 'Withholding Consumption: A Social Dilemma Perspective on Consumer Boycotts'. *Journal of Consumer Research*, vol. 28, no. 3, pp. 399–417.

Shachaf, P. and N. Hara (2010): 'Beyond vandalism: Wikipedia trolls'. *Journal of Information Science*, vol. 36, no. 3, pp. 357–370.

Svelch, J. and V. Stetka (2016): 'The coup that flopped: Facebook as a platform for emotional protest'. *First Monday*, vol. 21, no. 1.

Zaman, T. R., R. Herbrich, J. Van Gael, and D. Stern, 'Predicting information spreading in twitter'.

Boulard, C.; Castellani, S.; Colombino, T.; Grasso, A. (2019): Assessing the Intent and Effectiveness of Carbon Footprint Calculators. In: Proceedings of the 17th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing an the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep12

Assessing the Intent and Effectiveness of Carbon Footprint Calculators

Cécile Boulard, Stefania Castellani, Tommaso Colombino, Antonietta Grasso
NAVER LABS Europe
cecile.boulard@naverlabs.com

Abstract. In the context of addressing global warming issues, one of the possible approaches is to provide individuals with tools that support change toward greener practices, as for example around commuting. This paper illustrates a study that we conducted on the effectiveness of self-tracking of commuting data where participants received daily feedback on the financial costs and CO_2 emissions associated to their mobility practices. In the results, we describe situations where users do not accept the data and the models utilized to represent them, highlighting a limitation that diary instruments (and underlying models) of this type would have in supporting people to question and possibly change their mobility choices. On the basis of the study findings, we also describe a new model aimed at overcoming some of the limitations that the study showed, in particular by better connecting the individual environmental impact with the collective one.

Introduction

Global warming is a topic that raises many concerns at all levels in society. In response to these concerns the HCI research community has been involved in looking for solutions, especially in the area of limiting the impact of human activities on the environment (Bates *et al.,* 2018; DiSalvo *et al.,* 2010; Knowles *et al.,* 2018; Silberman *et al.,* 2014) and promoting change of practices to become

more sustainable (Froehlich *et al.,* 2009; Froehlich *et al.,* 2010). While we are aware that the Sustainable HCI community has raised questions about the effectiveness of addressing the global warming problem through individual level actions (Csutora 2012; Knowles *et al.,* 2018), it is also the case that there is a portion of the population who shows a willingness to better understand their personal footprint in order to engage in concrete actions to reduce and limit their environmental impact (Dolnicar *et al.*, 2008; McKercher *et al.*, 2010). In parallel, while sparse and country dependent, we are also aware of public initiatives directed toward sensitizing the population around environmental topics and providing incentives to change them.

In this context, in past work we started to study what role work organizations may play in being facilitators of awareness and change in domains like mobility and specifically commuting (Castellani *et al.*, 2014; Castellani *et al.*, 2016). More recently, we organized a follow up study, presented in this paper, where we wanted to address a group of users interested in assessing and possibly adjusting their commuting impact through the use of travel diaries. We setup a study with two main objectives.

The first objective was to further refine our understanding of how people commute using different modes of transport, how they choose among those, what the reasons are behind their choices, and what are the perceived advantages and constraints associated to each means. Our expectations were that having a better understanding of the decision-making process could inform the design of a more successful tool to incentivize commuters to be more conscious about their mobility practices. The second objective of the study was to test the usefulness and legibility of a standard model to track and measure financial costs and CO_2 emissions related to commuting, and to understand what would be the impact of using a travel diary in the decision process and in support of change behavior practices.

We targeted specifically financial costs and CO_2 emissions due to commuting as these quantities are at the same time hard to compute accurately for people and can have a relative strong impact on decision making for taking one or another means of transport. In order to be able to compare the understanding with the perception of the usage of several means of transport adopted for commuting, we focused the study on users that tended to use different means of transport for a similar trip. For participants using only the car or only the bus to go to work, the feedback in terms of costs and CO_2 emissions would have been the same every day automatically reducing the possibility for reflection offered by the self-tracking exercise.

Through the analysis of the interviews with the participants and their diaries we aimed to gather knowledge on how people reasoned about the different modes of transport and how they made their decisions. Moreover, we wanted to get some understanding of if and how the self-reflection on their commuting patterns may impact their future choice of means of transport, and also how each means of

transport was understood with respect to its environmental impact. What we did not really expect though was to have to face another aspect of the study, i.e. that a pretty standard way to compute the costs and CO_2 emissions associated to commuting would have been questioned in the way it was during the interviews.

This is aligned with what Remy and colleagues say, that there is more than usability to be evaluated when it comes to Sustainable HCI (Remy, 2018). Such a central aspect of a mobility self-tracking system became then the major focus of the subsequent analysis. It is the kind of serendipitous finding that may appear in a qualitative research and that in our case took over all other findings coming from the study (Corbin *et al.*, 2014; Rivoal and Salazar, 2013). These observations then led us to the design of a new model aimed at overcoming the limitations we found. In particular, we eventually proposed a new way to compute figures that participants of the study perceived to be both fairer, with respect to their impact on the environment, and more accurate.

In this paper we will focus on presenting the results that relate to how users dealt with the proposed model and the consequent design implications that we drew from the study.

Related Work and Study Objectives

The use of personal informatics, also referred as quantified-self or self-tracking, is today made possible by the variety of tools and connected objects that are available to individuals and has been widely analyzed in the HCI research community (Epstein *et al.*, 2015; Li *et al.*, 2010; Rooksby *et al.*, 2014). One recognized use of personal informatics is to support change management (Kefalidou *et al.*, 2015; Kersten-van Dijk *et al.*, 2017). The link between self-tracking and change management is in the reflexive position that users can adopt regarding their behaviors (Ptakauskaite *et al.*, 2018). Based on the collected data and on the change the user wants to achieve, personal informatics support the user by tracking progress toward a desired direction.

Many studies have focused on activity trackers that track the number of steps, the quality of sleep, the heart pulsation or burned calories. The majority of self-tracking practices target the domain of health and well-being (Choe *et al.*, 2014), where models are fairly simple: a step is a step, the quality of sleep computed as presence or absence of movements during the night, the number of heart pulsations, and the number of calories burned. Despite the required low level of knowledge and relative simplicity of the underlying models, some studies have already reported the difficulty that users may encounter when having to interpret the figures provided by the trackers (Coulter *et al.*, 2008; Herrmann *et al.*, 2018; Puussaar *et al.*, 2017).

When moving to the sustainable mobility domain, we can expect the difficulty of people in relating to the numbers to even increase, since the phenomena are much

more complex and difficult to reduce. If a user commutes with her car every day, how are we going to compute a fixed amount of money that is spent each day to go to work by driving a car? Should we include the initial cost of acquiring a car? The annual insurance fee? The costs of maintenance? If yes, how should it be integrated in the daily cost of driving a car to work? Then we should also add the cost associated to the fuel used for that specific trip, which is indeed the simplest thing to do when thinking of costs of commuting with a car. For the computation of CO2 emissions, in a similar way, many questions are open. How should we calculate the CO2 emissions of someone taking the bus? Should we take into account the number of passengers in the bus on that specific day? Or make an estimation with an average?

Current online tools for eco-feedback on mobility are based on disaggregated data among the various means of transport (in Ref. Carbon Footprint calculators). This means that for any mode of transport and a given distance, there is a cost and an amount of CO2 emission that is associated (for instance an average CO2 emission per distance and per passenger for public transport). The disaggregation of the data is a useful starting point as it has shown to help to support the understanding of the behaviors in settings like the smart grid (Froehlich *et al.*, 2011). However, as we will see through the results of our study, the use of disaggregated data coming straight from CO2 calculators is not enough to represent commuting in a way that users can relate to. As the computation is complex and is based on a range of factors (owning a car, having a public transport monthly pass, ride-sharing), it is difficult for a person to construct an accurate personal estimation of the impact of commuting practices (in terms of financial costs or CO2 emissions) (Betz *et al.*, 2010; Brazil and Caulfield, 2014; Waygood and Avineri, 2011).

This complexity translated by an absence of baseline and a possible under or over estimation regarding self-practices may lead to situations where a user may not understand or accept the data provided by the tracker. These difficulties, as we will see, may go beyond the difficulties highlighted in other studies of making data understandable through visualization and representation (Choe *et al.*, 2014 ; Rapp and Cena, 2016). The difficulties that we found are rather core to the definition of the model underlying the computation of costs and CO2 emissions due to commuting. There are studies on tools supporting greener mobility practices (Bie *et al.*, 2012; Bothos *et al.*, 2014; Bucher *et al.*, 2016; Gabrielli *et al.*, 2013; Jylhä *et al.*, 2013), but to the best of our knowledge, none of them details and discusses the model underlying the computation. This work aims at contributing to the body of knowledge about how people reason in practice about this type of data. Specifically, with this work, we want to provide the following contributions:

i) to highlight how, when it comes to self-tracking of abstract computed data, the choices of, what data to use, made by the underlying model impact the acceptability of the figures provided;

ii) to propose a more suited model for self-tracking of CO2 emissions and costs associated to commuting.

Methods and Settings

In order to study how people understand and accept figures of CO_2 emissions and costs due to commuting, a diary study has been conducted. We have chosen this methodology (Riemann, 1993) as it is difficult to gather information on commuting practices through observation over a long period of time. The first part of the study occurred over a period of 4 months in summer 2015 and the second part in 2016.
The study was undertaken in a French metropolis of 700 000 inhabitants which is a quite large city with characteristics in terms of transport infrastructure, mobility habits and needs of its inhabitants, etc., that may differ quite a lot with respect to a megalopolis or a small town. For example, in this town there is a well-developed public transport network available to its citizens (which is not necessarily the case for example in small towns). And this can have an influence on the way the participants to the study organized their commuting.

Participants

We recruited a group of participants from the city we are located in. It was important to be close to the participants in order to have a grounded understanding of the commuting context in the area, of the public transport options available, and to be able to facilitate the interviews during the study. The recruitment was done through a snowball sampling and ten participants took part in our research. This sample size is quite common in qualitative studies that targets to get a fine-grained understanding of a specific and complex question (Li *et al.*, 2012; Rapp and Cena, 2016; Thudt *et al.*, 2018). The ages varied from 27 to 56. Our participants were all professionals, qualified as engineers, computer scientists, doctors, technicians, school teachers and sales assistants.

The recruited participants all already had a commuting routine in place such that over a month they would use more than only one means of transport to go to work. Our objective with this constraint was to be able to provide participants with feedback on their behaviors looking at the different figures according to the various means of transport. The objective was to observe how participants would understand their environmental and financial impact related to commuting. This constraint was quite strong and made the recruitment process more difficult and longer than expected.

The home-work distance for the participants varied between 2.1km and 39.6km with an average of 15.1km (SD:13.9). The participants lived quite close to their workplace and our sample is consequently quite different from the typical

commuter average home-workplace distance in France that in 2004 was 25.9 km (Baccaïni *et al.*, 2007).
Table 1 shows the various means of transport used by the participants.

Table 1. Means of transport and Home-Work distance (km)

ID	Means of transport for commuting	Home-Work Distance (km)
P1	Tramway, bike, car	3.1
P2	Bike, bus, moped	6.7
P3	Tramway, car	3.4
P4	Bike, car	5.4
P5	Car-sharing, bus, car	27.6
P6	Bus, bike, car, car-sharing	18.5
P7	Bike, tramway, car	2.1
P8	Bike, train, car-sharing	39.6
P9	Car, car-sharing, bus	37.5
P10	Bike, bus, kick scooter, car	7.4

Procedure

The study was performed in two iterations. The first iteration is the diary study which was divided into three main steps. The first one was a face-to-face semi-directed interview with the participants where they described precisely their commuting practices: the means of transport used, the reasons for choosing a given means of transport on a given day, the preferences that they may have for one or another means of transport and the constraints they may have in their professional or personal lives according to commuting. The objective was to gain a global understanding of commuting practices. We also collected the exact path they used and all specific information on their personal car or moped (type, brand, year of construction, type of combustible). We also gathered information about possible goals associated to commuting, if they had any, such as: being able to read or do something else while commuting, reducing the financial cost of their commuting, increasing their physical activity, or limiting their environmental footprint.

During the second phase the participants filled a pen and paper diary during 20 working days. Each day, the participants had to indicate the date, the means of transport used that day, the reason why they made that choice and if anything pleasant or unpleasant occurred during the commuting. In the first interview, participants were asked if they would optionally share their diary with us every week in order to have weekly feedback of their practices. Even though many participants expressed enthusiasm regarding the proposal, only two of them exploited this possibility to receive weekly feedback. This relates to the extra work associated to collecting and managing the data (Lazar *et al.*, 2015). Nevertheless,

all of them finished the study and produced a daily diary as requested. After the 20 days, the participants returned their diaries and we were then able to compute all the figures regarding the cost and CO_2 emissions of their commuting. The model underlying the computation is the Eco-calculator described in the next section.

The third part of the study was a second interview with the participants, where we provided them with a compacted view of their commuting practices over the 20 days, as presented in Figure 1. For each day (a cell in the table), we represent for the user the means of transport used to go to work (top left corner of the cell) and the means of transport to go back home (bottom right corner of the cell).

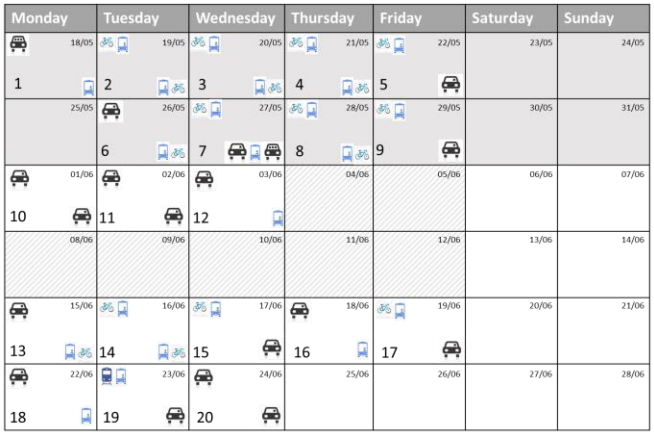


Figure 1 Compact representation of the commuting practices of P6 over the 20 days of diary study

We also shared with the participants a representation of the financial costs and CO_2 emissions for each day. Figure 2 shows an example of these representations for participant P6. All the figures produced on the participant's commuting was presented to the participants during interviews with all the required explanations in order to allow them to make sense of it.

The aim of that last interview was to assess if the exercise of keeping a diary on commuting practices had impacted the understanding of their practices and if any change had happened. The other goal of the second interview was to provide the participants with the compiled figures and then to discuss with them if and how those figures had made them think about their commuting practices and choices.

The second iteration used the exact same data from the participant's diaries to provide new figures of the financial costs and CO_2 emissions based on a new model designed to overcome the difficulties identified in the findings. The new model is also described in the findings. In order to question the new figures obtained and gather the feedback of the participants, we organized another round of interviews, only with the participants whose figures were impacted by the model. In our case it was all the participants who had use public transport during their 20 commuting days.

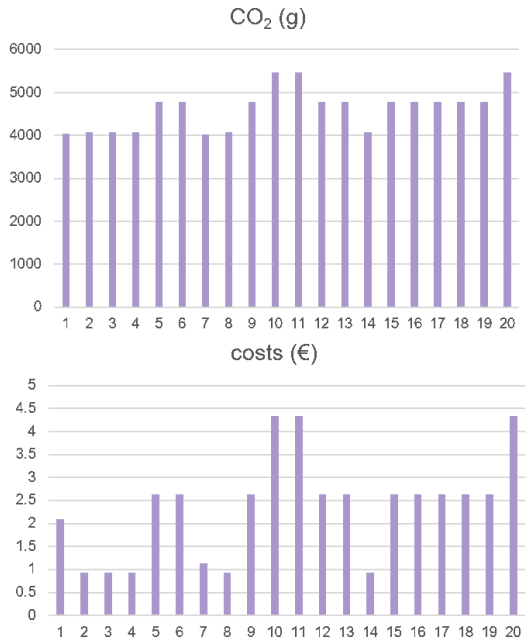


Figure 2. Feedback on the cost and CO2 emissions from the eco-calculator model due to commuting for P6

The eco-calculator

The computation of the financial cost and CO_2 emission figures was manually performed and based on typical carbon footprint calculators publicly available at the time of the study (in Ref., Carbon Footprint Calculator). The details are presented below.

For the trips with cars or moped: the cost estimation only considers the fuel consumption for the trip, meaning that for each travel, the computation is the distance covered (km) times the cost of a liter of fuel (€/L) times the average fuel consumption for that specific car or moped (L/km).

For the CO_2 emission, the computation is the distance covered (km) times the average CO_2 emission for that specific car or moped (g/km).

If participants did car-sharing, the figures computed for the cost and CO_2 emission were divided by the number of passengers for the travel.

For the trips with public means of transport, the cost is either the cost of a ticket or the monthly pass and the CO_2 emission is computed as the average CO_2 emission

for one passenger (g/km)[1] times the distance covered (km). For commuting using bike, kick scooter or walking, the costs and the CO_2 emissions are zero.

Finally, if the participant payed for a monthly transit pass, its cost was distributed over the 20 working days of the month.

Data Collection and Analysis

We audio-recorded all interviews and collected participants' diary entries. All the interviews were entirely transcribed and analyzed together with the diaries. We identified themes using a thematic analysis (Braun and Clarke, 2006). For this study, we inductively identified themes starting from the data trying to find commonalities rather than having a pre-existing representation of understanding. We describe our findings in next section.

Findings

Feedback from the interviews revealed situations where the data put the user at unease with respect to their own practices. Part of it relates to issues of understanding data and other aspects are rather linked to the acceptance of data. We do not want to develop this type of feedback in this paper, but rather to focus on a major outcome of the study which is the inadequacy of the intent behind the model, here to promote greener means of transport, and the figures provided to the participants.

Inadequacy between the model and its intent

The main objective of simulating a tool tracking commuting practices with a diary was to get a first evaluation of what could be its role, if any, in incentivizing users to use greener means of transport. It is therefore critical that the data is not only understandable, as discussed in the previous section, but also perceived as fair. Tracking data about CO2 emissions singles individuals and families out and questions their habits (and potentially their privileges) vis-à-vis a global problem. But to do so effectively, the model has to be able to properly contextualize the behavior of individuals and family units within the overall environmental impact of the collective (the city or metropolitan area, the country, etc.).

P5 lives more than 25 km from work and is used to commute either with a combination of car-sharing and bus, or by car solo, or by car-sharing. Because all options involve the use of the car, P5 and his partner were really concerned by their environmental footprint:

---

1 These figures come from the Carbon Footprint Calculator (In Ref. Carbon Footprint Calculator, first website) with the figures at the time of the study (2015) being for one passenger: 103,3g/km of CO_2 for the bus and 3.1g/km of CO_2 for the tramway.

"The main objective was to reduce costs, in terms of car mileage, petrol consumption, etc. Because we are both committed to being more environmentally conscious, and we said to ourselves that it amounts to having two cars less on the road, we decided to take the bus" (**P5**)

The solution they had chosen (Figure 3) was to have a main option with car-sharing from home to the bus-stop and then taking a bus. This meant that on the way back they had to coordinate to take the bus at the same time and then go back home together with the car. They had a second option, that was to go to work and back home with a car doing car-sharing. They resorted to this second option when one of them planned to use the car for some specific needs during the day.

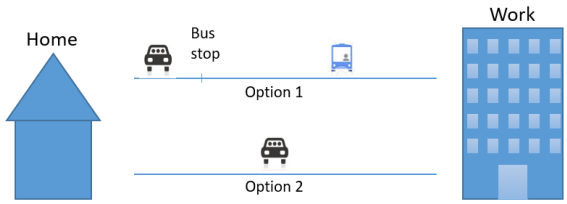


Figure 3. Options of commuting for P5

According to the eco-calculator model used in the study, the greener option was option 2. The calculator was showing that there were less CO2 emissions with option 2 than with option 1. When the researcher explained that to P5, he said:

"ah this is disappointing" (**P5**)

then he tried to understand by guessing that emission should have been a more global one:

"This is just for me, but in the bus, we are not alone" (**P5**)

The researcher explained again how the emissions of CO2 for public transport were computed (each passenger on a bus emits 103,3 g of CO2 per km). However, after the explanation, the questioning was still in place:

"which would mean that car-sharing would produce less CO2 than a bus even when it is full?" (**P5**)

This result was really surprising for P5 because it did not comply with the reality of facts as he perceived them: when P5 chose option 1, the car was used only on a little portion of the trip and then he took the bus. When P5 chose option 2 the car was used for the whole trip and there was also the bus circulating. So, from his perspective, option 1 cannot correspond to a higher emission of CO2 than for option 2 because there is one less car on the road between the bus stop and the workplace.

Similarly, P9 was surprised:

"Ah… I emitted less CO_2 in [car-sharing than in bus]" (**P9**)

Three other participants (P2, P6, P10) faced similar situations where for instance a trip using car-sharing or moped was causing a lower or almost equivalent in

quantity CO2 emission than a trip using public transport. This was not the representation of a "greener means of transport" according to what the participant had in mind.

To go further, P10 added:

"yeah the thing about the bus is that you are not responsible for the itinerary, you use something that is there regardless and you do not directly emit anything."

(**P10**)

Which expressed the gap between a global representation this participant had in opposition to a model that was providing daily feedback at an individual level.

Combining individual and shared responsibility in a new model

As mentioned in the introduction to this paper, one of the potential issues with this type of tool is that it is contentious whether that of individual behavior is the level where environmental questions can be effectively addressed. But when it comes to vehicle related CO_2 emissions, whether individuals decide to use public transport or greener forms of transportation, and on the basis of what information, is a matter of public policy and concern. And the fact is that the type of confusion over feedback provided by eco-calculators that we encountered in this study is particularly unhelpful when it comes to properly contextualizing individual behavior within a practice (commuting) that depends on public infrastructure and resources and is deeply impacted by choices made about them (for example, where and how to develop public transport, what incentives or disincentives are provided for the use of cars or other means of transport, public policies impacting cost of ownership, etc).

In order to provide feedback that more clearly contextualizes individual measurements and choices on costs and CO2, emissions within a public transport network, we have defined a model that explicitly considers and illustrates both urban community and individual related costs, and for the latter it encapsulates the different types of costs. The objectives were to increase accuracy regarding the real financial costs and CO2 emissions and fairness regarding the way people think about their commuting practices, and finally to have a model that can help to encourage the adoption of more sustainable means of transport. The model considers three types of costs (or CO_2 emissions): the community fixed costs, the individual fixed costs, and the individual variable costs.

Community fixed costs (CFC)

Community fixed costs can be computed mostly by using information provided by some public documents produced by urban area governments. Two types of information are required: information about the population and information about transport spending. For the first we decided to take as a reference the whole population living in the area, typically known from census data.

We did not make any distinction among commuters on the basis of their activities or home location as all of them can be at any point a user of public transport. For information about transport spending, we considered all public money spending associated to transportation. This included both road/infrastructure work and public transit sponsorship. For the CO_2 emission we considered only the information on the public transit. From these two figures we simply divided the total cost per number of inhabitants and per day in the year to obtain the community fixed costs.

Individual fixed costs (IFC)

Individual fixed costs are associated with two elements: the ownership of vehicles and the ownership of monthly or annual transit or parking passes. Ownership of a vehicle includes the cost of its acquisition, maintenance, and the insurance fee. For the figures used in the study, we collected all the required information listed for community fixed costs and individual fixed costs from a public document (SMTC, 2013) which is mandatorily produced by each French urban area larger than 100 thousand inhabitants. We have not included in our model any IFC CO_2 emissions but we could consider the CO_2 emissions of the production of a car or a bike.

Individual variable costs (IVC)

Individual variable costs are the costs that are generated when travelling in addition to the individual fixed costs. For financial costs these include the individual payments that are done when using a transport service, transit, taxi, parking car/bike rental, and the cost of the fuel consumed when using a private car. For CO_2 costs: the use of a private vehicle (including a taxi) accounts for the whole vehicle CO_2 emissions if used in single occupancy mode and divided by the number of occupants otherwise. Usage of transit services does not account for extra CO_2 emission since those are already included in the community fixed costs.

Figure 4 shows the new versions of the figures for participant P6 computed using the new model. When comparing with Figure 2 showing the figures for P6 with the first model, we directly see for days 2, 3, and 4 that P6 is using transit as she has no additional CO_2 emissions besides the CFC ones and there is no additional cost besides the CFC and IFC ones.

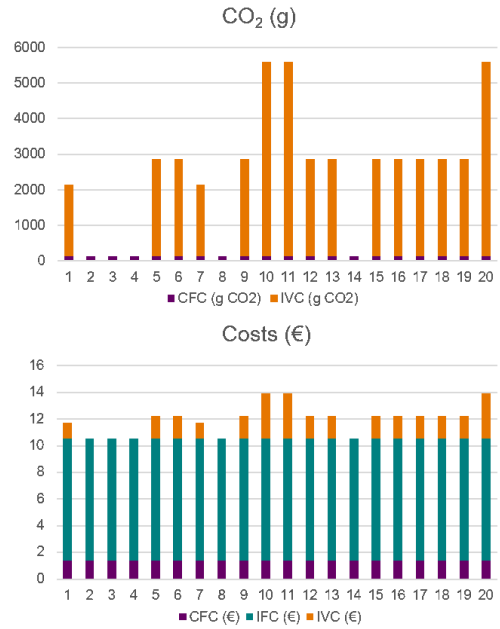


Figure 4. Feedback on the cost and CO2 emissions from the new model associated to commuting for participant P6

Feedback on the new model

The three types of costs illustrated in the new model can be applied to compute both financial costs and CO_2 emissions. We re-computed the data of all participants with the new model (as the second iteration of the study described in Material and Methods) and decided to conduct interviews using the new figures with the participants for whom it made a real change from the figures obtained with the first eco-calculator model (P2, P3, P5, P6 and P10). These were in fact the participants who used public transports during the data collection. All the interviews were entirely transcribed. During the interviews we were able to gather positive feedback on the new model. P2, P5 and P6 found that the new model was better at accounting for pollution issues. P2 and P3 identified that the costs computation was fairer and that it was relevant to have such notions. For P5 and P6, it was better incentivizing usage of public transit.

As a general conclusion, all the participants involved in this iteration with the new model found it to be more relevant, more accurate, fairer, and more convincing than the previous one.

We observed that the new model better represented the impact of user's choices and the possible impact of changes. This was obtained because costs and CO_2 emissions were organized showing both short-term and long-term impact of the user's practices or habits. The fact that there is public transport in a city depends

on the policies of a town and can be changed eventually by participating to elections. The cost of having a car is based on the decision of a user, at one point in time, to buy a car. This decision has a financial impact every day. Finally, the CO2 emissions due to a specific trip with a car is based on the choice on a given day to take the car instead of riding a bike, taking the bus or walking. In the end, it appeared meaningful to differentiate the data, not according to the means of transport (as in the eco-calculator), but rather according to the type of type of choice made by the user, where (s)he can indeed act.

Discussion

The intent behind the model

A main reason why the eco-calculator model was leading to inconsistencies was that the outcome of the calculation appeared to contradict the intent of the model itself (Lockton *et al.,* 2016). The intent of the eco-calculator was to support the adoption of greener practices. What appeared as an outcome was that, for several participants, the use of the bus led to more CO_2 emissions than car-sharing or moped. As P5 said:

"In that case I should stop taking the bus." (**P5**)

This conclusion would be very likely in opposition to what public authorities and common sense perceive as green transportation practices. Either the model inadequately computes the CO_2 emissions for feedback at an individual level, or car-sharing is really less polluting than public transportation and in that case it would be worth to acknowledge that and act accordingly.

In a similar fashion, we observed during the interviews side-effects from the new computing model. The intent behind the improved commuting model was both to avoid the sources of misunderstanding that we identified when using the first model and to provide figures showing to people the lower impact of some means of transport, like public transport or walking, cycling, or kick scooters. As this model makes visible the daily cost of ownership of a car, it might support users to more clearly consider the opportunity of owning a car or not. An unwanted side-effect of this kind of model is well described by P2:

"that is what I should tell myself, even when I take the bicycle I pay for the insurance of the car so it is not profitable to use my bicycle" (**P2**)

Behind any model there is an intent and possible unwanted side-effects.

Conclusion

The main outcome of this work, is that a model better tailored to provide feedback on costs and CO_2 emissions in comparison to currently used models, allows users

to get feedback that contextualizes their behaviors. It is beneficial to represent in the figures the relationship collective and individual responsibilities when it comes to commuting. Indeed, commuting requires to make choices upon individual options of means of transports, abilities, preferences in the context of a community that offers infrastructure such as public transport, cycle paths, roads etc. What the participants appreciated in the new model is that even though there is a relative high level of abstraction, this model was able to capture the complexity of the question and to reallocate the various levels of responsibility to make it fairer.

We believe that this work and the resulting new model can be inspiring for the quantified-self community about ways to answer to the need to better contextualize tracked data (Boulard-Masson *et al.*, 2018).

References

Baccaïni, B., Sémécurbe, F., & Thomas, G. (2007). Les déplacements domicile-travail amplifiés par la périurbanisation. *INSEE première*, *1129*(4).

Bates, O., Thomas, V., Remy, C., Nathan, L. P., Mann, S., & Friday, A. (2018, April). The future of HCI and Sustainability: Championing Environmental and Social Justice. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (p. SIG01). ACM.

Betz, M., Schwartz, T., & Ramirez, L. (2010). Know Thyself: Monitoring and Reflecting Energy Consumption. In *Workshop Know Thyself Extended Abstracts of the 2010 CHI Conference.* ACM

Bie, J., Bijlsma, M., Broll, G., Cao, H., Hjalmarsson, A., Hodgson, F., Holleis, P., van Houten, Y., Koolwaaij, J., Kusumastuti, D., and Luther, M. (2012). Move better with tripzoom. In *International journal on advances in life sciences,* vol.4, issue 3&4..

Bothos, E., Mentzas, G., Prost, S., Schrammel, J., & Röderer, K. (2014). Watch your Emissions: Persuasive Strategies and Choice Architecture for Sustainable Decisions in Urban Mobility. *PsychNology Journal*, *12*(3).

Boulard-Masson, C., Colombino, T., & Grasso, A. (2018, August). Analysis of 'Quantified-Self Technologies': An Explanation of Failure. In *Congress of the International Ergonomics Association* (pp. 579-583). Springer, Cham.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101.

Brazil, W., & Caulfield, B. (2014). Testing individuals' ability to compare emissions from public transport and driving trips. *Journal of Public Transportation*, *17*(2), 2.

Bucher, D., Cellina, F., Mangili, F., Raubal, M., Rudel, R., Rizzoli, A. E., & Elabed, O. (2016). Exploiting fitness apps for sustainable mobility-challenges deploying the goeco! app. *ICT for sustainability (ICT4S)*.

Carbon Footprint Calculators: https://www.transilien.com/fr/page-editoriale/le-calcul-des-emissions-de-co2, https://www.carbonfootprint.com/calculator.aspx

Castellani, S., Grasso, A., Willamowski, J., & Martin, D. (2014). Sustainable Commuting @ Work. *EAI Endorsed Transactions on Ambient Systems*, *1*(4), 1-5.

Castellani, S., Colombino, T., Grasso, A., & Mazzega, M. (2016, September). Understanding commuting to accompany work organisations' and employees' behaviour change. In *Smart Cities Conference (ISC2), 2016 IEEE International* (pp. 1-6). IEEE.

Choe, E. K., Lee, N. B., Lee, B., Pratt, W., & Kientz, J. A. (2014, April). Understanding quantified-selfers' practices in collecting and exploring personal data. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (pp. 1143-1152). ACM.

Corbin, J., Strauss, A., & Strauss, A. L. (2014). *Basics of qualitative research*. sage.

Coulter, A., Clegg, S., Lyons, G., Chatterton, T., & Musselwhite, C. (2008). Exploring public attitudes to person carbon dioxide emissions information.

Csutora, M. (2012). One more awareness gap? The behaviour–impact gap problem. *Journal of consumer policy*, *35*(1), 145-163.

DiSalvo, C., Sengers, P., & Brynjarsdóttir, H. (2010, April). Mapping the landscape of sustainable HCI. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1975-1984). ACM.

Dolnicar, S., Crouch, G. I., & Long, P. (2008). Environment-friendly tourists: What do we really know about them?. *Journal of Sustainable Tourism*, *16*(2), 197-210.

Epstein, D. A., Ping, A., Fogarty, J., & Munson, S. A. (2015, September). A lived informatics model of personal informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 731-742). ACM.

Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B., & Landay, J. A. (2009, April). UbiGreen: investigating a mobile tool for tracking and supporting green transportation habits. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1043-1052). ACM.

Froehlich, J., Findlater, L., & Landay, J. (2010, April). The design of eco-feedback technology. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1999-2008). ACM.

Froehlich, J., Larson, E., Gupta, S., Cohn, G., Reynolds, M., & Patel, S. (2011). Disaggregated end-use energy sensing for the smart grid. *IEEE Pervasive Computing*, *10*(1), 28-39.

Gabrielli, S., Maimone, R., Forbes, P., & Wells, S. (2013). Exploring change strategies for sustainable urban mobility. In *Designing social media for change at the ACM SIG-CHI conference on human factors in computing systems (CHI 2013)*.

Herrmann, M. R., Brumby, D. P., Oreszczyn, T., & Gilbert, X. M. (2018). Does data visualization affect users' understanding of electricity consumption?. *Building Research & Information*, *46*(3), 238-250.

Jylhä, A., Nurmi, P., Sirén, M., Hemminki, S., & Jacucci, G. (2013, September). Matkahupi: a persuasive mobile application for sustainable mobility. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication* (pp. 227-230). ACM.

Kefalidou, G., Skatova, A., Shipp, V., & Bedwell, B. (2015, August). The Role of Self-Reflection in Sustainability. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (pp. 1030-1033). ACM.

Kersten-van Dijk, E. T., Westerink, J. H., Beute, F., & IJsselsteijn, W. A. (2017). Personal informatics, self-insight, and behavior change: A critical review of current literature. *Human–Computer Interaction*, *32*(5-6), 268-296.

Knowles, B., Bates, O., & Håkansson, M. (2018, April). This Changes Sustainable HCI. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (p. 471). ACM.

Lazar, A., Koehler, C., Tanenbaum, J., & Nguyen, D. H. (2015, September). Why we use and abandon smart devices. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 635-646). ACM.

Li, I., Forlizzi, J., & Dey, A. (2010, April). Know thyself: monitoring and reflecting on facets of one's life. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems* (pp. 4489-4492). ACM.

Li, I., Dey, A. K., & Forlizzi, J. (2012). Using context to reveal factors that affect physical activity. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *19*(1), 7.

Lockton, D., Harrison, D., & Stanton, N. A. 2016. Design with intent. O'Reilly Media.

McKercher, B., Prideaux, B., Cheung, C., & Law, R. (2010). Achieving voluntary reductions in the carbon footprint of tourism and climate change. *Journal of sustainable tourism*, *18*(3), 297-317.

Ptakauskaite, N., Cox, A. L., & Berthouze, N. (2018, April). Knowing What You're Doing or Knowing what to do: How Stress Management Apps Support Reflection and Behaviour Change. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (p. LBW599). ACM.

Puussaar, A., Clear, A. K., & Wright, P. (2017, May). Enhancing Personal Informatics Through Social Sensemaking. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 6936-6942). ACM.

Rapp, A., & Cena, F. (2016). Personal informatics for everyday life: How users without prior self-tracking experience engage with personal data. *International Journal of Human-Computer Studies*, *94*, 1-17.

Remy, C., Bates, O., Dix, A., Thomas, V., Hazas, M., Friday, A., & Huang, E. M. (2018, April). Evaluation beyond Usability: Validating Sustainable HCI Research. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (p. 216). ACM.

Rieman, J. (1993, May). The diary study: a workplace-oriented research tool to guide laboratory efforts. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (pp. 321-326). ACM.

Rivoal, I., & Salazar, N. B. (2013). Contemporary ethnographic practice and the value of serendipity. *Social Anthropology*, *21*(2), 178-185.

Rooksby, J., Rost, M., Morrison, A., & Chalmers, M. C. (2014, April). Personal tracking as lived informatics. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (pp. 1163-1172). ACM.

Silberman, M., Nathan, L., Knowles, B., Bendor, R., Clear, A., Håkansson, M., ... & Mankoff, J. (2014). Next steps for sustainable HCI. *interactions*, *21*(5), 66-69.

SMTC. 2013. “2010 Le compte déplacements de l’agglomération grenobloise”. Retrieved June 26, 2018 from: http://www.adtc-grenoble.org/IMG/pdf/2013_Compte_deplacements_SMTC.pdf

Thudt, A., Hinrichs, U., Huron, S., & Carpendale, S. (2018, April). Self-reflection and personal physicalization construction. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (p. 154). ACM.

Waygood, E. O. D., & Avineri, E. (2011). Does" 500g of CO2 for a five mile trip" mean anything?

Towards more effective presentation of CO2 information. In *Proceedings of the Transportation Research Board 90th Annual Meeting* (pp. 23-27).

Hansson, K.; Sveningsson, M.; Sandgren, M. Ganetz, H. (2019): "We passed the trust on": Strategies for security in #MeToo activism in Sweden. In: Proceedings of the 17th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing an the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep14

"We passed the trust on": Strategies for security in #MeToo activism in Sweden

Karin Hansson[1], Malin Sveningsson[2], Maria Sandgren[3], Hillevi Ganetz[1]

[1] Stockholm University, [2] Gothenburg University, [3] Södertörn University

Corresponding author: khansson@dsv.su.se

Abstract. The #metoo movement can serve as a case for how networked online environments can provide settings for the mobilization of social movements, while also entail serious risks for those involved. In Sweden, over hundred thousand people were engaged in activities against sexual harassments and abuse, where social media were used to collect testimonies and to draft and discuss petitions that were later published in print news media. While HCI research on trust focus on how people trust technical systems, the authorities behind the system, or the user generated data, trust between peers in vulnerable communities is less researched. In this study, based on semi-structured interviews and a survey that involved 62 organizers of the Swedish #metoo movement, we therefore look into the question of how a secure and supportive environment was achieved among participants despite the scale of the activism. The result shows how trust was aggregated over networks of technical systems, institutions, people, shared values and practices. The organizers of the petitions used tools and channels at their disposal such as e.g. already established social media contexts that enabled the #metoo petitions to be formed easily and spread quickly. Establishing a supportive culture based on recognition and shared values was central for the movement. However, when the activism was scaled up, strategies were used to increase security by clarifying rules and roles, limiting access to information, restricting access to groups, and limiting the scope of communication.

Introduction

Networked online environments can effectively serve as settings for the organization and mobilization of social movements, e.g the environmental movement early on used social media to engage a broad public around substantive issues (DeLuca et al., 2016; Goodwin & Jasper, 2014; Pang & Law, 2017). Other examples of activism where social media played a role include the Arab Spring (AlSayyad & Guvenc, 2015; Smidi & Shahin, 2017), the Occupy Movement (Kavada, 2015), and movements such as the Gezi protests in Turkey (Haciyakupoglu & Zhang, 2015), and Ukraine's Euromaidan Uprising (Bohdanova, 2014).

Campaigns such as #metoo show how online spaces provide opportunities for victims of discrimination, harassment and abuse to come out and get support from other victims, and to participate in public debates around these issues. Simultaneously, research also points at the negative and practical consequences, which may render the digital feminist activism risky, exhausting and overwhelming (Mendes et al., 2018).

What characterized the #metoo movement in Sweden 2017-2018 is how well coordinated it was, despite being made up predominantly by grassroots initiatives. An important difference between the Swedish #metoo movement and other similar movements (such as the Arab spring) is that traditional media have usually been in opposition to the grassroots movements. In contrast, the Swedish #metoo movement used social and traditional media in a coordinated and remarkably efficient effort. The organizers managed to mobilize large groups through social networks, and then spread their agenda nationally through the largest and most influential newspapers. Judging from the public interest (Zachariasson, 2017), as well as the number of articles published in newspapers (Eklund, 2018), the Swedish #metoo movement can be described as very successful. The movement was also able to establish a feminist agenda focusing on structural problems, beyond the individual cases (Svärd, 2017). A broad mobilization took place in the form of lists of demands petitioned to the government, action plans by politicians and employers, as well as a large number of seminars and education organized around the country (Annebäck, 2018; Berglund, 2017; Samordningsgruppen för metoo, 2018). However, the framing of the movement as a success story obscures questions of obstacles that evolved along the way, concerning for example risks for those involved.

In the Swedish #metoo movement, perceived risks concerned not becoming employed, or losing one’s current employment because of the participation, or facing the social stigma of being a victim of sexual abuse. There were also fears of becoming target of threats or continued harassment. These risks had to be realized and handled in order for organizers to be able to gather participants and collect their stories and signatures.

This paper seeks to understand how the organizers of the #metoo petitions handled these risks and how they established a secure environment and gained participants' trust.

Background: Research on trust

In research on human computer interaction (HCI), trust is a central concept, as HCI to a large extent is about making people rely on the technology to solve different tasks. When navigating the topic of "trust online", the literature is dominated by research mainly on different types of e-commerce solutions (Corritore et al., 2003; Kracher et al., 2005), there is also research on e-government systems (Bannister & Connolly, 2011; Corbett & Le Dantec, 2018a, 2018b), and e-health systems (Beldad et al., 2010). The focus in these areas is mainly on how consumers and citizens can feel confident in systems that handle sensitive data such as money or medical records (Wang & Emurian, 2005). When it comes to trust in people, the focus has often been on the relationship between the citizen/consumer and the authority/service, and thus not directly about the trust in peers (Corbett, 2018a). A focus that is more about trust in peers is about trusting the reliability of user-generated data. The large amount of information available online creates an information crisis where trust in informal networks, rather than central institutions, are becoming increasingly important. For example, it may be in a situation where activists do not trust the official information, such as during the Gezi protests in Turkey (Haciyakupoglu, 2015). Here, the technology instead created an opportunity to "aggregate trustworthiness"(Jessen & Jørgensen, 2011) from a large number of sources, where social trust and technical affordances interact (Haciyakupoglu, 2015).

Another relevant aspect of online trust is personal security. Within this area, the relation between the desire for self-exposure and the possibility of being anonymous has been demonstrated when for example; it applies to sensitive subjects (Birnholtz et al., 2015), vulnerable groups such as victims of sexual abuse (Andalibi et al., 2016), or women who miscarriage (Andalibi & Forte, 2018). At the same time, research on people's safety awareness on social media shows that even though there are concerns that sensitive information is coming out, one chooses to trust that it works, as the benefits of sharing experiences and getting support are perceived as so valuable that it outweighs the risks. This also applies to vulnerable groups such as illegal immigrants in the United States (Guberek et al., 2018).

Undoubtedly, trust is something central to online communication and also a broad and multifaceted concept that means different things in different contexts, why for the sake of clarity we here would like to define it and explain how we relate to the concept in this paper.

Following Haraway(1991), technology can be seen as a kind of prosthesis, which extends our "arms" and allows us to stretch beyond our bodies and reach what we previously could not reach. In this view, trust is about trusting that the arms can reach out and carry what we expect them to do. There is always a risk that the prosthesis will fall off, but most of the time it goes well. The moment of risk means that trust is required, which is why risk and trust are closely associated. The more risk, the greater is the trust needed.

When it comes to technologies such as social media, these are not primarily artifacts but consists of humans, sometimes very large amounts of people that one might not even have a personal relationship with, but it might be a common interest that brought one together. In these cases, the trust is not so much a matter of trust in technical systems, trust in authorities, trust in information, or trust in particular people, but trust in shared values and practices. For example, it may be about belonging to an idea, or a shared experience, which is sufficiently strong or revolutionary to motivate the individual to, for example, take the risk of trusting strangers in publics (Wang, 2005).

Trust is also linked to distance. Simplified, the greater the distance, the greater the trust required. It can be about physical distance, temporal distance, emotional or social distance (Corbett, 2018a). Here, trust can be seen as a process of bridging distances, a process that can be described in various phases such as *developing*, *building*, and *maintaining* trust (Rousseau et al., 1998). In the development phase, trust is about *a calculated and weak confidence*. Trust in this phase is mainly cognitive and is about relying on clear evidence and strong external structures such as laws and systems. In the construction phase, trust is more about *experience built through interactions over time*. People and situations that have previously been reliable are trusted again. The third phase, maintaining trust, is less about calculations and more about belonging, and takes its point of departure in *shared values and benevolence*. One not only trusts that the system will work, and that people are predictable, but one trusts that this is motivated by shared values.

Data and method

To understand how different factors such as social trust and technical affordances played a role in the organization of #metoo, this study employs a mixed methods approach, consisting of a survey and semi-structured interviews.

The survey was distributed to the organizers of all 79 petitions that were initiated between November 2017 and June 2018. The number of contact persons varied per petition group, as well as how contact information was provided. Some groups provided group-aliases that transferred e-mail to all the organizers of the petition in question, and others provided individual addresses of one or a few of

the organizers. Some petitions were organized by groups of people, whereas others by just one person. The petitions differed in reach as well: some gathered over 10 000 participants, whereas other groups were smaller and more closely knit. We distributed the survey to 105 organizers and got responses from 62 organizers of 50 petitions within two weeks.

The 62 organizers of the 50 petitions (see appendix 1) came from all over Sweden, from Malmö in the south to Kiruna in the very north. One person lived in the neighboring Finland. 31 lived in Stockholm, the capital of Sweden, and 4 in Gothenburg, which is the second largest city. The remaining 27 respondents lived in different small towns or rural areas. The organizers were between 20 and 70 years old with the majority (44 of 62) between 30 and 50 years old. Educational levels were high, 54 of 62 had a college education, which is twice as many as in the general Swedish population (SCB, 2018).

The seven interviewed informants ranged from being in their twenties to 50+. Their previous experiences as organizers were mixed, from no experience at all to a lifelong experience of media activism. Before becoming the organizer of a petition, many of them already had access to some sort of professional network online; they could for example serve as moderators for social media groups gathering people from their industry or be responsible for an e-mail lists that connected former classmates.

The survey asked questions about how the petitions were organized, what tools and methods were used, how news media was contacted, and what role security and trust played in the petitions' organization.

As a way to get complementary information and to deepen our understanding, semi-structured interviews were conducted with seven petition organizers, from a number of contexts: IT industry, construction industry, the forest industry, agriculture, and equestrian sports. Informed consent was gathered, and all names of participants and the petitions have been omitted to ensure anonymity.

Each interview lasted between 50 and 70 minutes and began with a brief overview of the purpose of the research, followed by a series of questions asking the informant; to describe their background and role in the organization of the petition, what ideas and values that influenced the organization, how the petition was organized, how it was distributed, about the role of security and trust, and what they had learned from the experience.

The interviews were recorded and transcribed. All data were in Swedish, thus the quotes have been translated into English. The interview material, as well as open-ended questions in the survey, were analyzed thematically in an approach inspired by grounded theory, where a first open-coding of the data was followed by more focused coding to develop salient categories. The paper focuses on the result of the survey of communication tools and processes and on one of the main themes that emerged from the analysis of the open survey answers and the interviews; security strategies.

Results

Communication tools and processes

The tools for initiating and developing the petitions varied from IRL meetings to phone calls, e-mail, collaborative writing to social media and survey tools. Social media, in most cases meant Facebook. Facebook was used by almost all petitions. E-mail or messenger were used in half of the cases, and Google docs was used in 1/3 of the petitions. Twitter and Instagram were foremost used in addition to Facebook, as a way to distribute the petitions.

After the initiating phase, testimonies were collected through e-mail, social media and survey tools. One fourth of the petitions used some sort of survey tool to collect testimonies and signatures, the other either used a Facebook group, or had a dedicated e-mail address. In the final distribution phase when the result of the petition was communicated to its stakeholders, social media was central, but also IRL meetings such as seminars, meetings with journalists and decision makers, became important.

The smaller group of organizers usually used a Facebook group, Messenger group or chat as an exclusive channel to communicate among themselves. Most organizing groups maintained a close and continual contact through different tools.

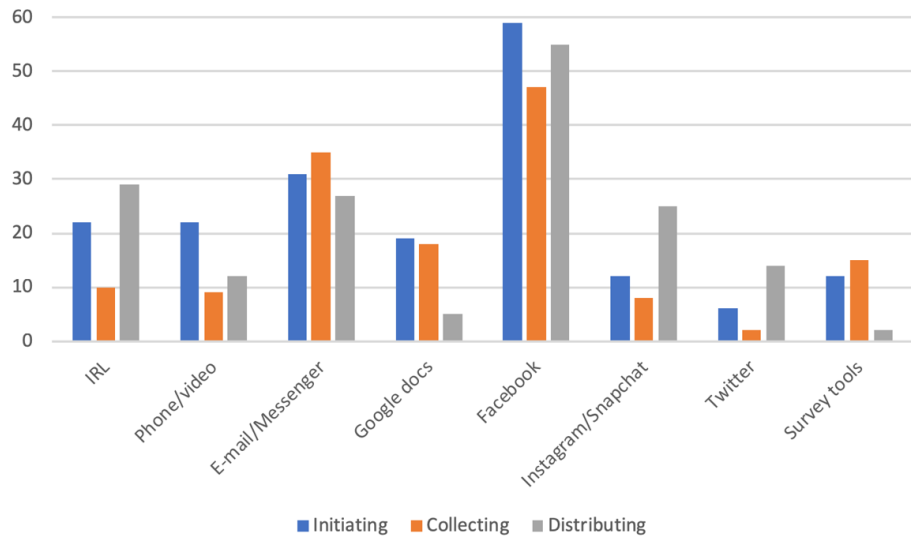


Figure 1. The amount of organisers that used different tools in different parts of the process; when *initiating* petitions, *collecting* testimonies, and *distributing* petitions.

The organizer of the petitions used a combination of methods and tools to communicate, and the processes could look very different from case to case. For example, one petition started as a discussion thread in a Facebook group that

already gathered thousands of women from the industry. When the first #metoo petitions were published, the issue was discussed in this discussion thread and several participants asked for a petition in their own industry. In response, one of the participants in the discussion thread quickly put together a manifesto and set up a survey tool to collect signatures online and created an e-mail address that others could send their testimonies to. She posted information about the petition in the open Facebook group, which meant that the petition quickly received a wide distribution. After a few hours she had enough material for drafting an article proposal aimed to a leading newspaper. The whole process from idea to draft went very fast and took no more than 4-5 hours. To get help to develop and complete the article, she asked the others in the discussion thread, and in this way a group was formed around the continued work.

Other petitions developed much slower. It was not uncommon that testimonies were published semi-public in closed Facebook groups, which generated long discussions on each individual case. In some petitions, the text was developed collectively, not only by the organizers, but all members of the group came with opinions, and the text was examined in detail and discussed intensively before it was sent for publication. Several such cases took place in Facebook groups with over thousands of participants.

The interviews show that later petitions learned from previous petitions' experiences and were thus more cautious about how they e.g. used social media or with publishing their private email addresses.

To sum up, the organizers of the petitions used tools and channels at their disposal such as already established social media contexts. Most often different tools and channels were combined. Facebook seems to have been used by almost all petitions and had a central role in the whole development of the petitions both as a way to reach out and as a forum for discussion.

Strategies for security

While the whole idea of #metoo was to make sexual harassment visible and defy the shame of having been exposed, it meant great risks for the victims to come out with their stories. The disclosures could, for example, lead to unpleasant consequences both socially and professionally, in the form of social exclusion, threats and harassment. Legally, accusations that cannot be substantiated in concrete evidence or other witnesses, can lead to the person reporting the case being sentenced and punished for defamation. On the other hand, perpetrators pointed out publicly may suffer from extreme consequences that are not proportional to the possible crime. It is therefore not surprising that one central themes in the open questions in the survey and in the interviews were about security strategies.

The material points to five comprehensive strategies for security where the first emphasizes a supportive community and openness, and the other four are strategies for security with the aim to regulate and control.

Security through a supportive community

Most petition organizers emphasized the importance of a trustful environment where the participants dared to talk about their experiences and could receive support and encouragement from others with similar experiences. The study shows how they gave special importance to a supportive culture in which the victims were not questioned and were a generous culture where established through active and collective moderating.

Establishing safe and trusting forms of dialogue was central to the organization of the petitions. It was crucial to create a situation where people who previously might never have told others about their experiences, could get recognition for these, and open themselves up without being questioned or risking their identity coming out. The situation was based on confidence in the organizers and their ability to harbor trust.

> "Most stories were submitted to me and [the other organizer]. Some released their stories in the Facebook group, which created trust so that other people also dared to share. This in turn created trust. But to send by e-mail felt more secure and we were careful to ask before posting the stories that this really was ok. I am thinking that we showed great respect, from the beginning, and that this was a good start. We as organizers set rules for what we could talk about and not in the group.
>
> I believe the security of these groups is largely based on the evidence of how widespread the problem is. If, for the first time, you feel that you are listened to and taken seriously, and if you feel for others in the group, then the interests to break the social rules is not so great." (Survey answer from organizer of one of the petitions)

Being recognized, transparency, and having seen everyone else contributing, created another kind of trust, a trust in the community of a collective experience. Contributing with a traumatic experience became meaningful when they got feedback from a large group and heard other share their experiences, and when they come out this contributes to more sharing. An encouraging environment for discussion characterized by generosity and without judging or blaming, was an explicit ambition that was raised by several organizers of the petitions. This supportive culture was also reproduced by the participants.

Several of the organisers also had a readiness to handle people who needed more support, for example by providing information about people or organizations that provide legal or psychological support.

> “We as administrators and coordinators took an active role and set the tone in the comment fields. There were never hatred or bullshit, instead many pointed out how good the mood was. The focus was on "Thank you for telling us" and always reminding you that there was the opportunity to get more support. We worked a lot with responsiveness and for example using languages that did not exclude. From the very beginning, we created an opportunity for anyone who wanted to talk to a person in charge at our federal office if they needed more

support and / or wanted to report a perpetrator, to possibly move on in some way." (Survey answer from organizer of one of the petitions)

The organizers were subjected to a lot of pressure during an intense time period, and in the questionnaire about how they did to create security and safety in the organizing group, we get many similar answers that emphasize a communicative strategy where all means are used to have close contact with the group:

"We had our own WhatsApp group where we supported and pepped each other all the time."

"A lot of conversations, support and a constant checking with each other."

"Mainly through active contact and support between us. We replaced each other when there were tougher discussions in our Facebook group."

"Continuous communication between us, we met a lot and talked a lot about what was sent to us."

"We who organized were in different places so could not meet physically but had close contact over Messenger so we would always be on the same wavelength."
(Survey answer from five organizers of different petitions)

Some organizers knew each other personally before, which facilitated communication. But many were not familiar before the call, or just superficially familiar before, and found each other through the shared engagement on the issue.

Security through clear rules and roles

A strategy that contributed to creating a trusting environment in many petition groups was the development and communication of clear rules. Attitudes were also developed and disseminated between the petition groups, largely via the overall coordination group, which gathered the organizers.

"Clear directives on publishing in the group. We were clear about how we safeguarded anonymity and total anonymization of testimonies (no one was allowed / could publish testimonies in the group. Testimonies were first sent to e-mail that we admins later published without names and places or other "disclosure" in the group). Additions to the group needed to be approved by the contact person and us in admin. The group was secret and not searchable." (Survey answer from organizer of one of the petitions)

Active moderation was another source of security. Those who moderated the groups worked actively to ensure that the rules on anonymization and generosity were complied with, and they closed down discussion threads that didn't follow the code of conduct. They also reminded the participants about the rules and the goal of the campaign as a way of improving the level of discussion.

Likewise, another safety measure was to appoint one or a few people who acted as spokesmen for the group. Clearly speaking for all the anonymous voices of the petition was a way of removing focus from individual organizers, and instead emphasizing a collective voice.

Security by limitation of information

The trust was also based on the fact that names did not spread - neither on victims nor perpetrators. Technical affordance was fundamental to effective

implementation. The technical security was (somewhat surprising) nothing that the organizers experienced as risky, instead it was the human factor that they could be worried about: it was crucial to be able to trust that members of the group did not spread the name and information further.

A basic principle that all groups have embraced was anonymity: The right to be anonymous, but also to let others, even perpetrators, be anonymous in the testimonies. Active moderators ensured that this was complied with in testimonies as well as group discussions in social media. This mainly concerns what is communicated externally and to other members, but in a few groups, there were full internal anonymity, ie those who left testimonies were anonymous also to the organizers, and that the organizers were anonymous to the participants.

All groups had restricted access to information about the victims, and to the uncensored testimonies. In the relatively open groups, however, many testimonies were published directly by victims, which meant that the person then became known to the whole group which could consist of thousands of people, sometimes with serious consequences.

> "To avoid testimonies leaking from the group, we started collecting them in a separate document and deleting them from the Facebook group. This turned out to be too late. A woman was contacted by her perpetrator after her testimony leaked." (Survey answer from organizer of one of the petitions)

Following this event, members were asked to send their testimonies either directly to the organizers, or through a form that allowed full anonymity. Here, different considerations needed to be taken into account and balanced against each other. While it was important that information did not leak out, the sharing of testimonies and feedback on these stories was important to develop a trustful atmosphere that made more people dare to testify. This was resolved in some petitions by making the administrators share the testimonies on social media, allowing the victims to be kept anonymous, while people still were being able to discuss the testimony and publish their support.

> "We had rigid rules on anonymity in the group, for having the security to share. This meant that it was mainly us administrators that shared the testimonies in the Facebook group." (Survey answer from organizer of one of the petitions)

It seems that the need to be anonymous was perceived as being particularly important in tightly connected networks, where everyone knew everyone. Firstly, because there was too much to lose if it came out that you participated in the #metoo activism, as there were few possible new workplaces to switch to. Partly, the perpetrators, or the perpetrators' relatives, were often well known and included in their social network, which meant that they (or other people with an interest in the issue, such as human resource managers at companies), easily could access information in social media by looking over the shoulders of a partner or simply by sharing login information with a family member.

Some petition organizer therefore chose to be completely anonymous and did not have any named organizers or signatures at all. This approach of total anonymity, even towards journalists, could create difficulties in reaching out and gaining legitimacy, but was sometimes a necessary way to go to avoid reprisals from colleagues and family, or for fear of what the public light would entail.

Security by limiting access

One way to ensure that information did not seep out of the group was to carefully check and limit new members. A related security issue concerns the power imbalance of the group, to ensure that unauthorized persons did not gain access to the group:

> ”Here is the crux ... it was decided, for example, that no journalists would be allowed to join the group. Then part of the admin group went in with the argument "but that's my friend" and added these people again. Also industry professionals / service persons at [industry name] were added with the same argument against the group's will, even though the group assumed to be a group for [professional identity]. Thus, they were expected to tell sensitive stories to their employers, who in some cases leaked information into [the industry company].” (Survey answer from organizer of one of the petitions)

Here they were careful about not accepting managers or employers as participants, or teachers mixed with students. A closely related dilemma is the cases where some member is closely related to a person who has power over the others in the group. Sometimes participants left the group voluntarily as they experienced that their participation reduced the trust within the group. But many times, questions about power imbalance and dependency conditions could be a dilemma, which sometimes made the organizers take other paths, and use other tools than just Facebook:

> “We didn't want to bring in some of them [who had high positions or worked as a person responsible for staff] who we knew about. So, we never arranged such a [Facebook group], but instead we spread the Google form via Messenger and yes, we sent it to our nearest network, and so it spread. So then it became so that one could pass it around and say that it comes from a safe source. It was as if we passed the trust on.” (Interview)

For most petitions, the question of who would participate was simple: Women in the industry. Many petitions spoke in the names of women and non-binary. But in several cases, discussions arose about the question of who would be allowed to participate. The least controversial was the separatism, to exclude non-women, as including men was seen as the presence of potential perpetrators and could reduce the trust within the group. But in industries where the career paths were a little unclear, a discussion also emerged about the significance of boundary drawing and why industry-specific manifestations had an importance.

> "Many people signed the petition, but many have a very vague connection to the industry, but more willingness to be seen and heard, and to be in the limelight that the [...] industry has. Whether someone harasses you in your amateur [context] says more about how society is

at large than how […] the industry looks. ” (Survey answer from organizer of one of the petitions)

Foremost, the issue with participation was not about professional identity but about power. Since sexual harassment is seen as an expression of a power structure where those who consider themselves to have more power are those who harass them with less power, it was an important issue that the participants in the group had a fairly equal status so that they were not in different ways potentially contributing to these power schemes, e.g as a manager or client.

The importance of other power structures such as age and sexuality was also a discussion that came up, and made the requirement for equality within the group complicated. The affinity with other vulnerable people collided with the affinity of colleagues, family or others of the same age or other forms of power positions.

"I took the initiative to a meeting irl afterwards, it was very strengthening to meet people, but I reacted on that most of the people who came to the meeting were heterosexual white women in their 40s-50s .... we did not recruit the young, perhaps because one of the members of the admin did not want to have students in [context] because she taught [there]. I thought we should have included the young. that is my opinion, because they are the weakest and perhaps the most vulnerable, at least it has been so historic. ”(Survey answer from organizer of one of the petitions)

Security by limitations in scope

Another security strategy applied by several petitions was limitations in scope. The gathering of testimonies and signatures could e.g. take place for a limited time and the group was then closed down when the petition was published. Another aspect of scope concerns the size of the group. Although it was seen as positive that the petition created interest and engagement, problems arose if they become too big.

The challenge of scaling up a feminist supportive culture developed in smaller groups was made clear by the speed of how the calls were developed, where quick decisions must be taken without any formal leadership. The larger the group became, the more uncertain it became for the participants, as the possibilities for information leaked increased. But above all, it was labor intensive to moderate large, sprawling group discussions that went on around the clock.

“Our security was never a problem. The most problematic was workload and stress. ” (Survey answer from organizer of one of the petitions)

It happened that groups grew uncontrollably, and some administrators felt that they did not manage to administer the group, and that they could not control that the information was not leaking.

“The larger the group became, the more unsafe it became. We tried to make those who wrote in the group aware of this and think about that we became bigger and that the secrecy became increasingly difficult to maintain. However, the stories that became public were anonymized and when it was published we closed down the group for reasons of confidentiality.” (Survey answer from organizer of one of the petitions)

One measure taken to reduce stress was to "pause" the group for a period, when it was not possible to post, to give all participants a much needed rest from the intensive discussions in the forums.

In summary, most groups started from a feeling of trust based on recognition and shared values. But especially when activism was scaled up, strategies were needed to increase security by clarifying rules and roles, limiting access to information, restricting access to groups, and limiting the scope of communication.

Levels of security

At an overall level, the groups applied similar strategies for security, but there was a varying level of security that could be divided into three type groups.

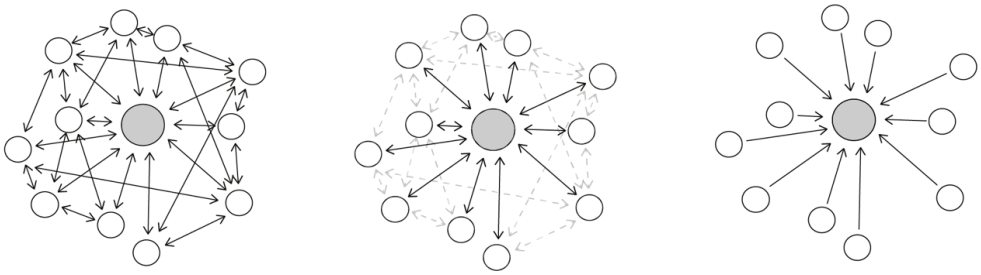


Figure 2: Illustration of information flow on three levels of security: *Community*, where information flow in all directions between the organizers and participants, and participants and participants, and security is based on trust on shared values and community; *Regulation,* where information flows in both directions between the organizers and participants but where the organizers acts as gatekeepers moderating the information flow between participants and participants; *Alienation,* where information flows from participants to organizers, but participants have no means to contact other participants, and can be anonymous also for the organizers.

Some groups applied a *light security* that relied on fellowship. Organization and collection of testimonies were made in closed groups on social media, to which people belonging to the defined group were invited. The invitations to the group worked according to the snowball principle so that everyone was invited by someone who knew them, so there was a social closeness and community between the participants of these expanded networks.

Other groups had *moderate security* where communications were more regulated. On Facebook, they created hidden and "secure" groups that weren't searchable, that sometimes changed names constantly so that they would be harder to find for outsiders. Anyone invited was checked by the group's administrator, who in some cases also moderated posts before they were posted to ensure that no one posted names or anything else that could harm the safety of

individuals. In this security strategy trust was based on common rules and leadership.

Some groups applied a *strict security* based on alienation where a (sometimes completely) anonymous organizational group collected testimonies via a survey tool. Those who participated had no opportunity to contact each other or see the other participants' contributions. Here, the technology was used to minimize the risk that identifying information were spread by minimizing the possibility for the participants to communicate with each other or with the organizers.

Within these different levels of security, various strategies were used to promote trust and ensure security: a supportive community, clear rules and roles, limiting information, limiting access and scope.

Concluding discussion

The results show that security was a central issue in the organization of the petitions, where various strategies were used to promote trust and ensure security: a supportive community, clear rules and roles, limiting information, limiting access and scope.

The safety of the individuals was important for trust in the organization of the call, but at the same time there was a contradiction between being anonymous and feeling a sense of community and trust in the collective. The organization of the petitions employed a varying degree of security, corresponding with the degree of perceived risk from the participants, ranging from an emphasize on *belonging* and relationships to focusing on *regulation*s, to a situation where the tools and methods enabled *alienation* as a mean to reduce risk:

- From a process of trust through shared values, benevolence and *belonging*. Here, the starting point was the trust due to social closeness and a shared interest with people from an enlarged network. Information flowed in all directions between the organizers and between participants and participants.
- To a process of trust based on pronounced *regulation* and leadership. Here you have learned to not trust anyone who wants to be involved: Information flow in all directions between the organizers and participants and participants, but the organizers acted as gatekeepers and censors.
- To a process of trust that involves calculating and strict security where confidence is weak and the technology is used to enhance *alienation* between users and thus minimized risk and create a social distance. Information flow from participants to organizers, but not the other way.

No organizers expressed concern regarding the risk that the technology would not work or any privacy concerns of technical nature. This tendency is confirmed by previous research, that people choose to rely on technology, even though they should know better (Guberek et al., 2018).

In trust research trust is seen as a process of crossing distances, a process that can be described in various phases such as developing, building, and maintaining trust (Rousseau, 1998), where one phase is based on the other and moves towards trust based on increased community and belonging. In the case of #metoo in Sweden, one can see the different types of trust processes as an expression of the levels of trust in the different industries from which the petitions originated. But one can also see it as a reverse trust process, where trust initially was high, in the beginning when the group was smaller, and then were reduced when the organizers realized the risks and when the groups became larger and exposed to the public. The biggest risk was also experienced in situations where the participants actually came from a tightly knit network and thus knew or understood each other well. Here the risk was that the participants could have conflicting loyalties.

Both the participants and the organizers initiated and participated in processes they seldom had control over and rarely had previous experiences of. The strength to actually implement these risky projects came from previous successful petitions that acted as role models and established a shared set of values and practical examples. Technical affordance was another important factor. A number of easily available technical tools functioned as prothesis that enabled the organization to be scaled up and extended to thousands of participants. Several petitions were made in horizontal networks in social media that organized women in the industry, and there were often already established networks that enabled the #metoo call to be formed easily and spread quickly. The trust that enabled so many to actually participate was not trust in a particular system, trust in an authority or a known person. Instead, the trust can be described as aggregated, it was established through trust in technical systems, institutions, people, shared values and practices, and lots of trust-generating interactions over time both before and during the actual organization of the petitions.

References

AlSayyad, N., & Guvenc, M. (2015). Virtual Uprisings: On the Interaction of New Social Media, Traditional Media Coverage and Urban Space during the 'Arab Spring.' *Urban Studies*, *52*(11), 2018–2034. doi:10.1177/0042098013505881

Andalibi, N., & Forte, A. (2018). Announcing Pregnancy Loss on Facebook: A Decision-Making Framework for Stigmatized Disclosures on Identified Social Network Sites. *In Proc. of CHI*, 1–14. doi:10.1145/3173574.3173732

Andalibi, N., Haimson, O. L., De Choudhury, M., & Forte, A. (2016). Understanding Social Media Disclosures of Sexual Abuse Through the Lenses of Support Seeking and Anonymity. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16* (pp. 3906–3918). doi:10.1145/2858036.2858096

Annebäck, K. (2018, January 24). Så går fackförbunden vidare efter #metoo. *ETC*. Stockholm.

Bannister, F., & Connolly, R. (2011). The Trouble with Transparency : A Critical View of Openness in e- Government. *Policy & Internet*, *3*(1), 158–187. doi:10.2202/1944-2866.1076

Beldad, A., De Jong, M., & Steehouder, M. (2010). How shall i trust the faceless and the intangible? A literature review on the antecedents of online trust. *Computers in Human Behavior*, *26*(5), 857–869. doi:10.1016/j.chb.2010.03.013

Berglund, B. (2017, March 8). Metoo – men sedan då? Det här hände i några av de utpekade branscherna. *Metro*.

Birnholtz, J., Merola, N. A. R., & Paul, A. (2015). "Is it Weird to Still Be a Virgin." In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15* (pp. 2613–2622). New York, New York, USA: ACM Press. doi:10.1145/2702123.2702410

Bohdanova, T. (2014). Unexpected Revolution: The Role of Social Media in Ukraine's Euromaidan Uprising. *European View*, *13*(1), 133–142. doi:10.1007/s12290-014-0296-4

Corbett, E., & Le Dantec, C. A. (2018a). Going the Distance: Trust Work for Citizen Participation. In *2018 CHI Conference on Human Factors in Computing Systems* (p. 312). doi:10.1145/3173574.3173886

Corbett, E., & Le Dantec, C. A. (2018b). Exploring Trust in Digital Civics. In *Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS '18* (pp. 9–20). doi:10.1145/3196709.3196715

Corritore, C. L., Kracher, B., & Wiedenbeck, S. (2003). On-line trust: Concepts, evolving themes, a model. *International Journal of Human Computer Studies*, *58*(6), 737–758. doi:10.1016/S1071-5819(03)00041-7

DeLuca, K. M., Brunner, E., & Sun, Y. (2016). Constructing Public Space|Weibo, WeChat, and the Transformative Events of Environmental Activism in China. *International Journal of Communication*, *10*(June 2014), 19. doi:321–339 1932–8036/20160005

Eklund, J. (2018, September 12). *#METOO - En studie om Metoos gestaltning i svenska medier*. Gothenburg University. Retrieved from https://gupea.ub.gu.se/handle/2077/57614

Goodwin, J., & Jasper, J. M. (2014). *The social movements reader : cases and concepts* (3. ed.). Oxford: Wiley-Blackwell. Retrieved from http://libris.kb.se/bib/16775789

Guberek, T., McDonald, A., Simioni, S., Mhaidli, A. H., Toyama, K., & Schaub, F. (2018). Keeping a low profile? Technology, risk and privacy among undocumented immigrants. *Conference on Human Factors in Computing Systems - Proceedings*, *2018–April*, 1–16. doi:10.1145/3173574.3173688

Haciyakupoglu, G., & Zhang, W. (2015). Social Media and Trust during the Gezi Protests in Turkey. *Journal of Computer-Mediated Communication*, *20*(4), 450–466. doi:10.1111/jcc4.12121

Haraway, D. J. (1991). *Simians, cyborgs, and women : the reinvention of nature*. Routledge: New York.

Jessen, J., & Jørgensen, A. H. (2011). Aggregated trustworthiness: Redefining online credibility through social validation. *First Monday*, *17*(1–2), 1–8.

Kavada, A. (2015). Creating the collective: social media, the Occupy Movement and its constitution as a collective actor. *Information Communication and Society*, *18*(8), 872–886. doi:10.1080/1369118X.2015.1043318

Kracher, B., Corritore, C. L., & Wiedenbeck, S. (2005). A foundation for understanding online trust in electronic commerce. *Journal of Information, Communication and Ethics in Society*, *3*(3), 131–141. doi:10.1108/14779960580000267

Mendes, K., Ringrose, J., & Keller, J. (2018). #MeToo and the promise and pitfalls of challenging rape culture through digital feminist activism. *European Journal of Women's Studies*, *25*(2), 236–246. doi:10.1177/1350506818765318

Pang, N., & Law, P. W. (2017). Retweeting #WorldEnvironmentDay: A study of content features and visual rhetoric in an environmental movement. *Computers in Human Behavior*, *69*, 54–61. doi:10.1016/J.CHB.2016.12.003

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not So Different After All: A Cross-Discipline View Of Trust. *Academy of Management Review*, *23*(3), 393–404. doi:10.5465/amr.1998.926617

Samordningsgruppen för metoo. (2018, March 5). 65 metoo-upprop: "Här är vallöftena som vi vill se" - DN.SE. *Dagens Nyheter*. Stockholm, Sweden.

SCB. (2018). Utbildningsnivån i Sverige [Educational level in Sweden]. *SCB*. Retrieved January 16, 2019, from https://www.scb.se/hitta-statistik/sverige-i-siffror/utbildning-jobb-och-pengar/utbildningsnivan-i-sverige/

Smidi, A., & Shahin, S. (2017). Social Media and Social Mobilisation in the Middle East: A Survey of Research on the Arab Spring. *India Quarterly: A Journal of International Affairs*, *73*(2), 196–209. doi:10.1177/0974928417700798

Svärd, C. (2017). *"Två korta ord: me too#" En studie över #metoo:s nyhetsvärde och gestaltning i den svenska dagspressen*. Gothenburg University. Retrieved from https://expo.jmg.gu.se/wp-content/uploads/1027.pdf

Wang, Y. D., & Emurian, H. H. (2005). An overview of online trust: Concepts, elements, and implications. *Computers in Human Behavior*, *21*(1), 105–125. doi:10.1016/j.chb.2003.11.008

Zachariasson, H. (2017). 7 av 10 svenskar tror att me too-uppropen kommer leda till förändring [7 of 10 Swedes believe the metoo petitions will lead to change]. Sweden: SVT Nyheter.

Appendix 1

Table: List of the 50 hashtags that where included in the survey; translation, date, publisher, amount of signatures, and context.

Hashtag	Translation	Date	Publisher	Signatures	Context
#tystnadtagning	Silence, camera action	2017-11-10	Svenska Dagbladet	705	Actors
#imaktenskorridorer	In the corridors of power	2017-11-17	Svenska Dagbladet	1 319	Politicians
#närmusikentystnar	When the music becomes silent	2017-11-17	Dagens Nyheter	1 993	Music industry
#tekniskfel	Technical problem	2017-11-19	Dagens Nyheter	1 139	Technology industry
#deadline	Deadline	2017-11-21	Feministiskt perspektiv	4 084	Journalists
#inteförhandlingsbart	Not negotiable	2017-11-22	ETC	1 501	Union movement
#tystdansa	Silent dancing	2017-11-22	Feministiskt perspektiv, Dagens Nyheter	620	Dancers
#akademiuppropet	Academics' petition	2017-11-24	Svenska Dagbladet	2 400	Academics
#omniberättarlyssnarvi	If you tell we will listen	2017-11-24	Dagens Nyheter	1 299	Psychologists telling about clients' experiences
#vardeljus	Let there be light	2017-11-24	Kyrkans tidning	1 382	Swedish church
#sistaspikenikistan	Last nail in the coffin	2017-11-27	Aftonbladet	4 672	Construction industry, architects
#vårdensomsvek	The health care that failed	2017-11-27	SVT Nyheter	none	Patients
#larmetgår	The alarm is on	2017-11-28	Aftonbladet	none	Emergency services
#sistabriefen	Last brief	2017-11-28	Dagens Nyheter	2 126	Communications industry
#givaktochbitihop	Stand firm and suck it up	2017-11-29	Dagens Nyheter	1 768	Swedish defense
#nykterfrizon	Sober free zone	2017-11-29	Accent - Sveriges största tidning om droger och nyhkterhet	500	Recovered alcoholics
#visparkarbakut	We are bucking	2017-11-29	Dagens Nyheter	1 089	Equestrian
#skiljaagnarnafrånvetet	To separate the wheat from the chaff	2017-11-30	ATL Lantbrukets affärstidning	937	The Green industry
#utantystnadsplikt	Without professional secrecy	2017-11-30	Svenska Dagbladet	10 400	Physicians
#utgrävningpågår	Excavation running	2017-11-30	Dagens Nyheter	387	Archeologists
#vikokaröver	We are boiling with rage	2017-11-30	Dagens Nyheter	1 863	Restaurant industry
#metoobackstage	Metoo backstage	2017-12-01	Svenska Dagbladet	1 614	Television, film and stage production
#slådövörattill, #byss	To turn a deaf ear	2017-12-03	Dagens Nyheter	634	Deaf community
#ålandocksåå, #högtskalldetklinga	Åland also	2017-12-03	Egen hemsida	1 568	Finnish citizens living on Åland
#konstnärligfrihet	Artistic freedom	2017-12-05	Konstnärernas riksorganisation	1625	Arts and crafts
#nomore	No more	2017-12-06	Dagens Arena	none	School management
#lättaankar	Weigh anchor	2017-12-07	Sjöfartstidningen	1000	Shipping
#utanskyddsnät	Without safety net	2017-12-10	Dagens Nyheter	none	Persons in addiction, criminality or prostitution
#virivermurarna	We are tearing the walls down	2017-12-10	Aftonbladet	954	Prison and probition service
#skrattetihalsen	Choking the laugher	2017-12-11	Dagens Nyheter	80	Comedians
#intedinhora	Not your whore	2017-12-16	Dagens Nyheter	144	Persons in prostitution
#nustickerdettill	Now it will hurt	2017-12-16	Dagens Nyheter	1 309	Health care employees

#bortabrahemmavärst	There is no place worse like home	2018-01-02	SVT Nyheter	718	Persons with experience of domestic violence
#dammenbrister	The pond is breaking	2018-01-02	Astra	6111	Finno-Swedish citizens
#inationensintresse	In the interest of the nation	2018-01-03	Uppsala Nya tidning	826	Students in Uppsala
#obekvämarbetstid	Uncomfortable working hours	2018-01-10	Handelsnytt	377	Commercial employees
#inteminskuld, #påvåravillkor	Not my debt, on our terms	2018-01-14	Dagens Industri	300	Banking, financial and insurance industries
#nostranger	No stranger	2018-01-18	Expressen	500	Victims of racism
#slutvillkorat	No more conditions	2018-01-22	Feministiskt Initiativ	60	Persons with normbreaking disabilities
#nödvärn	Self-defense	2018-01-25	Nödvärn	5 000	The police
#allmänhandling	Public document			Not published	Governmental employees
#exponerad	Exposed			Not published	Photographers
#fordonsindustriupproppet	The transport industry petition			Not published	Transport industry
#husfrid	Domestic peace			Not published	Against domestic violence
#ikulturarbetarnasrum	In the room of the cultural workers			Not published	Cultural workers
#kidstoo	Kids too			Not published	Association for persons related to sexually abused children
#vispelarintemed	We do not play along			Not published	Game industry

Sunnen, P.; Arend, B.; Heuser, S.; Afkari, H.; Maquil, V. (2019): Designing Collaborative Scenarios on Tangible Tabletop Interfaces – Insights from the Implementation of Paper Prototypes in the Context of a Multidisciplinary Design Workshop. In: Proceedings of the 17th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing and the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_ep13

Designing collaborative scenarios on tangible tabletop interfaces - insights from the implementation of paper prototypes in the context of a multidisciplinary design workshop

Patrick Sunnen[1], Béatrice Arend[1], Svenja Heuser[1], Hoorieh Afkari[2], Valérie Maquil[2]
[1]University of Luxembourg, [2]Luxembourg Institute of Science and Technology
Corresponding author: patrick.sunnen@uni.lu

Abstract. Within the context of the research project ORBIT (Overcoming Breakdowns in Teams with Interactive Tabletops), we design and study a joint problem-solving activity at an interactive tabletop, that gives participants the opportunity to develop their collaboration methods. To gain design insights for the development of a scenario soliciting participants to collaborate, we set up a multidisciplinary design workshop. During the latter, we explored and discussed three different collaborative scenarios, implemented as paper prototypes. In this paper, we report on first results gained from an exploratory analysis of the video data that was recorded in the context of this workshop.

Introduction

Shared interfaces such as multi-touch tables and tangible tabletop interfaces were repeatedly found to mediate and support collaboration. Ioannou and Antoniou (2016) summarize that tabletops enhance the sense of teamwork, sollicit interaction and willingness to participate in group tasks, increase equity in physical interaction

and facilitate joint attention on the task. These benefits are largely due to the big shared screen and the possibility for direct and simultaneous interaction by multiple users (Mercier and Higgins, 2014). When participants' attention is drawn to the tabletop, they can see each other's actions as well as the system's feedback, potentially changing the nature of the collaboration (Price, 2013). So, explicit awareness of other's (hand) actions can facilitate explorative conduct and increase collaborative forms of construction and interpretation (ibid.).

While multi-touch tabletop interfaces are operated using finger touches, tangible tabletop interfaces (TTI) additionally make use of physical objects that can be placed, moved or rotated in order to interact with the system. Due to their physical nature, a TTI can be conveniently embedded in a real physical space and situated in a social setting (Fernaeus et al., 2008). In particular, the physical objects support participants in partitioning and coordinating their activities (Scott and Carpendale, 2004), and facilitate individual ownership and announcement of tool use as support for group awareness (Speelpenning et al., 2011).

A vast body of research has already identified how the design of TTI enables multiple users to jointly work on a shared task or enhances group work (Fleck et al., 2009; Yuill et al., 2012; Stanton et al. 2001; Woodward et al., 2018). Our work contributes to and attempts to extend these previous works in two aspects. First, we focus on a specific understanding of collaborative conduct. Second, we seek to create and identify design aspects which go beyond 'just' enabling participants to collaborate, but furthermore elicit them to collaborate.

In everyday life and in some literature, the term 'collaboration' is often used very broadly to describe two or more persons working together on the same task. However, in our work, we go beyond this general understanding of collaboration and to do so, we mainly rely on Roschelle and Teasley (1995). They define collaboration as a coordinated, synchronous activity where mutually engaged participants rely on a mediational framework to construct and maintain a negotiated and *shared emerging conceptual space* to jointly solve a problem (according to a shared understanding of the latter). The above-mentioned conceptual space is referred to by the same authors as "Joint Problem Space (JPS)" to grasp how collaborative activity gets organized in participants' interactions. JPS incorporates participants' orientation to (shared) goals, their descriptions of the current problem state, their awareness of available problem-solving actions, and associations interrelating the previous aspects. So, the JPS is considered here as an interactional achievement rather than as a convergence of individuals' mental representations (Sarmiento-Klapper, 2009).

Within the context of the ORBIT-project (Sunnen et al., 2018), we design and study a joint problem-solving activity at an interactive tabletop, that gives participants the opportunity to develop their collaboration methods. To develop design implications for that matter, more precisely, for the development of a scenario soliciting participants to collaborate, we set up a multidisciplinary design

workshop. During the latter, we explored and discussed three different collaborative scenarios, implemented as paper prototypes. In this paper, we report on first results gained mainly from an exploratory analysis of the video data that was recorded in the context of this multidisciplinary design workshop.

Designing collaborative scenarios

So, the very first design question that arises from the above described perspective on collaboration is the following: How can we design a TTI-mediated joint problem-solving activity that elicits collaborative conduct among the participants, or in other words, the construction and maintenance of a JPS? More specifically, we focus here on the design of the TTI, which is meant to be a fundamental component of the mediational framework through which participants establish and maintain a joint problem space.

Thus, in the course of developing the design, we retained the following 'preconditions of collaboration': the TTI is supposed to afford the co-construction of a shared semiotic space as well as to solicit and sustain the participants interactions as mutually organized. In line with these prerequisites, three intertwined TTI aspects can be varied in order to explore their consequences on the collaborative conduct of the participants: the difficulty of the task, the complementary distribution of participants' competencies[1] and the organization of the physical semiotic space of the TTI.

A group-worthy *task* should be challenging and equally addressing all the participants, and invite them to work together interdependently and reciprocally to reach a common goal. This can be achieved if the task aims at creating a situation in which participants' exchange of ideas and information, and their joint construction of understanding are vital to success (Mercer, 1995; Vass and Littleton, 2010; Cohen and Lotan, 2014). Within each of our scenarios, we rendered the sub-tasks more and more challenging by adding further constraints at each level.

Closely related to the task are the competencies that are assigned to and realized by the participants. In order to have the possibility to participate in the accomplishment of the task in a mutually engaged way, participants need to be provided with complementary abilities so that they have to rely on multiple resources that cannot be mobilized by one person alone. The complementarity of the competencies was implemented here as a mobilization of tangibles in time, either simultaneously or sequentially ordered. Note, that even though the TTI-activity pre-determines what can be done and what cannot not be done, meeting the

[1] By 'competency' we mean here the potential abilities and roles 'provided by' or 'built into' the TTI-mediated joint activity. Whether and how these competencies are actually embraced and enacted by the participants is, of course, a different story and constitutes a primary concern of our analytical work.

task-challenge via competencies ‘is not given’ but has to be, explored, negotiated and (or at least) coordinated by the participants.

We already mentioned that the large shared screen of a TTI is of paramount importance when it comes to supporting the construction of a joint focus among the participants. So, we decided to explore the physical semiotic space of the interface with regard to directionality and visible access, and in terms of the organization of the space (parcels, fields, connected space).

Thus, the following, more specific intertwined design questions emerged for us:

1) How can we organize the physical space of the TTI to solicit the construction of a joint conceptual space?
2) How can we design TTI-instantiated complementary competencies so that they elicit participants' mutual engagement with one another to construct and maintain a negotiated and shared emerging conceptual space?
3) How can we design a challenging task that solicits participants mutual engagement in a joint problem-solving activity?

We then tailored these three aspects to our context, goal and target audiences and the outcome turned out as three scenarios (see table 1), which we tested as paper prototypes during our multidisciplinary design workshop. In the following, we shall give more information on the design workshop, the three scenarios and how we evaluated the latter.

Multidisciplinary Design Workshop

A central element in all of the scenarios was a shared central space, where all the participants have equal access to the current state of the game. All three versions were designed to be 'played' by three adult participants with no required training or specific skills. After defining the details of each scenario such as the main goal, tasks, roles, levels and challenges, we made a paper prototype of each game to test them in the design workshop. The scenarios were developed by a team of two computer scientists. The latter also participated as moderators in the workshop, and a team of three social scientists[2] participated as users (without being aware of the exact game mechanics). The social scientists, furthermore, provided a feedback from the perspective of researchers investigating collaborative conduct. The aim was to evaluate the aspects of collaboration in each scenario and decide about the features to consider for further development. The session lasted in total four hours and was audio and video recorded. The participants played each scenario on average for 30 minutes and there was on average 40 minutes of discussion after each test session.

[2] The involved computer and social scientists are also the authors of this paper.

Table I. Overview of the designed and evaluated collaborative scenarios

	Task	**Complementary competencies**	**Organization of space**
Scenario 1: Damaged spaceship (Figure 1)	Retrieving various specified parts (appearing randomly in the different parcels) with the fitting tools	Retrieval and carrying means are distributed among the three participants: • every participant can carry 2-3 parts, • every participant can use his/her two exchangeable tools to retrieve a part placed in one of the three terrains, • later, an extra tool is needed to get the parts (two participants must simultaneously mobilize tools).	Three enclosed, rectangular parcels representing different terrains (desert, ocean, forest).
Scenario 2: Growing crops (Figure 2)	Cultivating various types of crops on the fields by applying different farming resources in a specific sequence	Farming resources and seeds are distributed among participants: • participant has tractor and wheat seeds, • participant has water and bean seeds • participant has fertilizer and orange seeds Sequence: tractor, seed, water, fertilizer	Eight closed areas with different shapes and sizes (from 1 to 6 units) representing fields to be cultivated.
Scenario 3: Collecting garbage in the see (Figure 3)	Steering a ship to specific positions in the open sea to collect items (garbage and later fuel) and to return ship to harbor	Movement options distributed among the three participants: • participant in the North (N) can move southward (S) and southeast (SE) • participant in the Est (E) can move westward (W) and northwest (NW) • participant in the South (S) can move northern (N) and northeast (NE) Movement to E is only possible through an alternation of NE and SE. Movement to SW is only possible through an alternation of W and S Movement to SW is only possible through an alternation of W and S.	One connected space representing the sea with several islands and a harbor.

Each test session started with the explanation of the 'game' by the computer scientists (as moderators), followed by the pilot level to let the participant familiarize themselves with the features of each scenario. Then, the participants played different levels of each scenario with one of the computer scientists acting the reactions of the computer, moving and placing the objects of the paper prototype. During each test session and discussion, all the members (testers and moderators) were taking notes of the remarks and the raised ideas. At the end of each session, the participants discussed the experience, focusing on the potential of the scenario to trigger collaborative conduct as well as the suitability of the scenario to be instantiated in various contexts. After the workshop, we went through the recorded materials to further investigate the scenarios from the perspective of collaborative conduct. The latter is what we report on in this paper.

Description of the three scenarios

Scenario 1: Damaged spaceship

Figure 1. Picture of the damaged spaceship scenario.

For the first scenario (Figure 1), we subdivided the central space of the tabletop into three parcels representing different terrains: ocean, forest and desert. Participants were told that they were astronauts and had to repair their spaceship.

To achieve this goal, they had to collect different parts (provided as picture cards on the tabletop), which were scattered over the three different terrains. To collect the parts, specific tools (provided as picture cards[3]) were needed and each terrain required participants to use two different tools (for example, a hammer and a torch) for pick-up. After being informed about the required number and type of the different parts (visualized through a list), they needed to have the right tools to retrieve the required parts, which appeared randomly in the different terrains. Every participant had a personal area with two slots, where the previously collected tools could be placed. To use the latter, participants had to touch the part with their avatar (an astronaut). The part was then moved to the designated place disposing of 2 or 3 slots (according to the level) in the personal area of the collecting participant.

These constraints were here our way of implementing complementary competencies among the participants. Everyone could only store two respectively three parts and dispose of two tools. So, the main collaborative task in this scenario was therefore distributing the tools among them to collect the parts. We expected the participants to discuss their strategies and to coordinate their actions with regard to picking up the right part at the right time. In the last level, to emphasize the coordination challenge of the task, three tools were needed to fetch a part[4]. To do so two of the participants had to simultaneously touch the part with their avatars.

Scenario 2: Growing crops

The second scenario was set in a farming context (Figure 2). The shared space was divided into eight areas of different shapes and sizes designated as fields. Every participant received a tangible representing a bag of seeds (wheat, bean, orange) and a tangible providing him/her with the control over a farming resource (tractor, water, fertilizer). The set goal of this scenario was to grow certain amounts of the available crops in different fields.

To reach this goal the participants had first to discover and then apply the procedure to cultivate a field. As soon as a tangible is placed on the shared space, participants receive a feedback (green check or red cross) from the TTI whether the tangible was applied at the right moment in the sequence (which is: tractor, seed, water and then fertilizer). Therefore, they had to try out and explore together different combinations of using their competencies to make the products grow. To keep the task challenging and to further solicit discussions and coordination efforts among the participants, a number of constraints were introduced along the levels (adjacent fields cannot contain the same product, amounts to grow are given, time constraint).

---

[3] In the TTI implementation the parts would be provided as digital objects and the tools as tangibles.

[4] Two tools related to the terrain and one extra tool related to the part.

Participants' main collaborative task was here to coordinate their actions to apply their complementary competencies in the right sequence to cultivate a field successfully. Furthermore, they had to discuss and agree on cultivating strategies (where to plant, what to plant and how much). The task was considered as accomplished when participants had harvested the asked amount and so the overall success was the result of the joint performance of all participants.

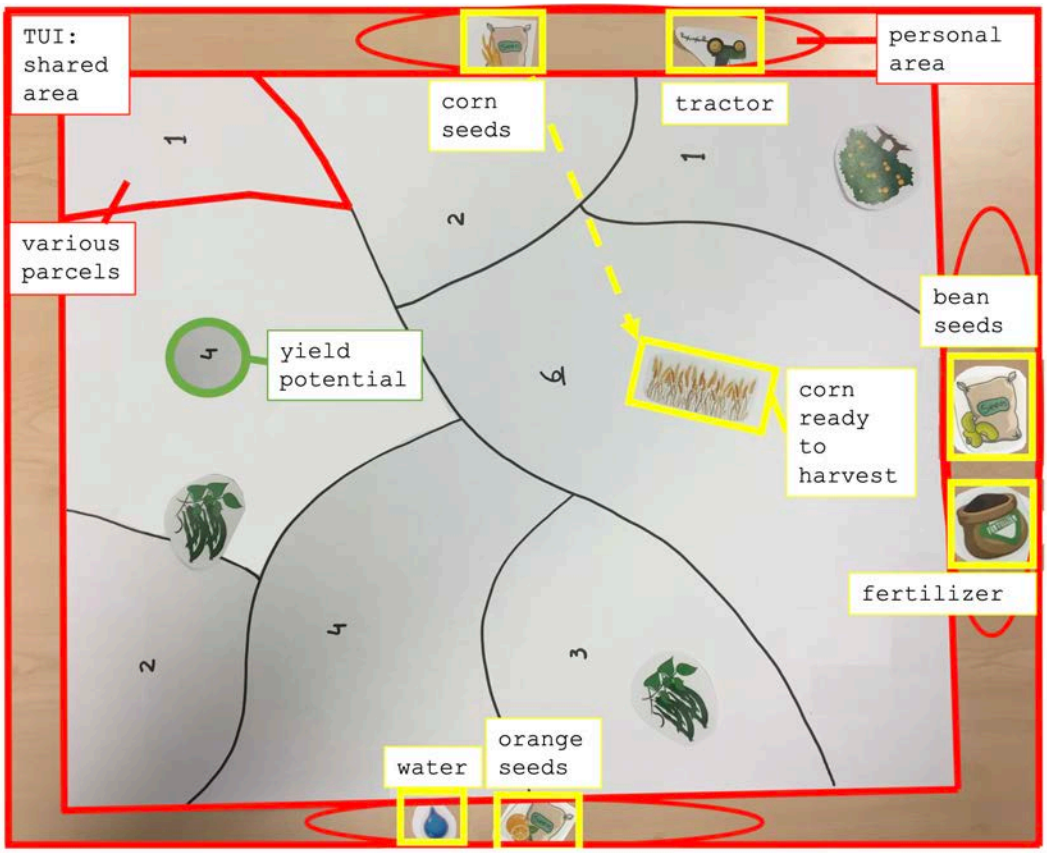


Figure 2. Picture of the growing crops scenario.

Scenario 3: collecting garbage in the sea

The third scenario was inspired by Piper et al. (2006) and the central space consisted of an 8*8 grid representing an ocean with some islands and a harbor (Figure 3). As common goal participants were asked to collect with their ship a certain amount of randomly distributed garbage items, while avoiding crashing into an island. At the end of each level they had to return their ship to the harbor. Reaching this goal became more challenging in later levels, since we introduced fuel usage (1 unit per movement, restorable through refills) and time constraints (limited availability of garbage items, overall time limit).

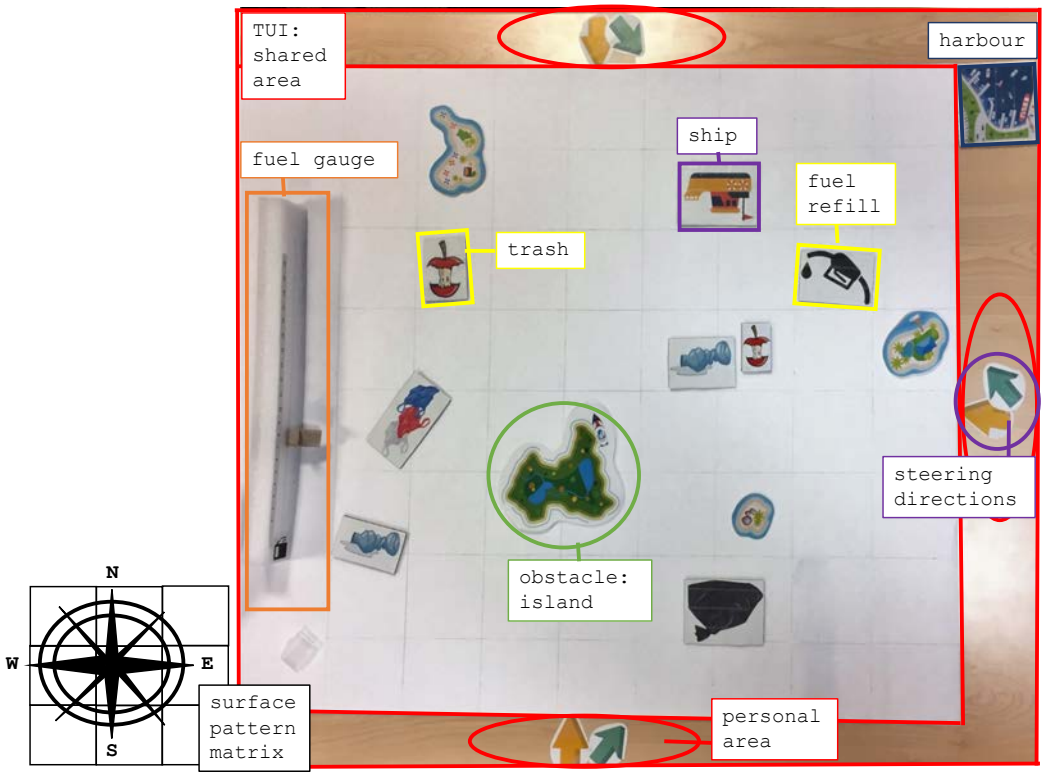


Figure 3. Picture of collecting garbage scenario[5].

The collaborative challenge of the task was that participants had to steer the ship together to reach a targeted location since the movement options (their competencies) were distributed among them in a complementary way. Indeed, each person was given the ability to steer the ship in just two different directions by taping on one of the two arrows situated in his/her personal area. The ship would then move by one cell per tap in the required direction.

The resulting consequences of these movement options (see Table 1) were the following: First, only six directions were immediately available; second, two directions could only be taken via the composition of two other directions (allocated to two different persons); and, third, the chosen route could only be taken by sequentially operating the different - distributed - steering widgets. Consequently, in order to successfully accomplish the task[6], participants had to agree (ideally after a mutually engaging discussion) on the items to target as well as on the route to take, and they had to coordinate their steering actions.

[5] The wind rose and the surface matrix (on the left side of the picture) are depicted here for the convenience of the reader and were not part of the design.

[6] Collecting the required amount of garbage items and returning to the harbor (levels 1-4), without running out of fuel (levels 2-4) or time (levels 3-4) and without crashing into an island.

Exploratory analysis of the three scenarios

Scenario 1: Damaged spaceship

The batch of three parcels constituted the central space of the TTI and all the participants had visual access to the three regions. Perhaps unsurprisingly, this parceled organization of a large part of the tabletop space did not elicit the construction of a joint focus in the same way all along. After discussing the task of the respective level, participants organized the allocation of the tools (competencies) in such a way, that they had at their disposal the requested pair of tools providing them with the ability to collect the parts located in the terrain closest to them. The terrains being exchanged (by the moderator) after each completed level, this interactional work was achieved several times. To get this distribution done (see transcript of extract 1 as an exemplary instance), they were mutually engaged (all three participants participated equally in the exchange), oriented to a shared goal (solving the task efficiently by allocating the terrains to participants), described the current problem state (e.g., lines 6 and 8) and were aware of problem-solving actions (e.g., lines 1, 10 and 12). So, they constructed a shared understanding of the problem and established a joint conceptual space (JPS).

Transcript of extract 1 (17:16-17:53)

01	P^1	we could negotiate and say ((...)) ah Patrick (.)you could focus on (.) that ((pointing at list of collectable parts))
02	P^2	we have to be careful
03	P^1	and you ((pointing at P3)) can focus on that ((pointing at list of collectable parts)) and I could
04	P^3	maybe we should focus on the (.) the worlds ((tapping at each parcel))
05	P^1	or on the worlds (.) yes
06	P^2	because the problem if you focus on this ((pointing at list of collectable parts))
07	P^1	yea (.) yea
08	P^2	you will not have the right tools
09	P^1	yea
10	P^3	or you need to say oh ((pointing at P^1)) please pick it up now ((pointing at ocean parcel))
11	P^1	Yes
12	P^2	What we could do (.) we could exchange the tools and everybody is closer to his territory because now my territory is there ((pointing at desert parcel))

However, once this allocation negotiations were concluded, the shared focus became less discernible as a visual instantiation. Overall, participants tended to focus more on their terrain, and waited for the requested parts to appear and retrieved them (Figure 4). After the completion of level 1, one participant made this explicitly accountable by saying to the moderator "I was focused on that (pointing to forest parcel) because I had these (pointing to her tools) (...) so that was mine (laughing)" (Figure 5). Notice that during the activity the participants categorized one another with labels such as "forest lady" or "desert space man" thus emphasizing the previously established connection between a participant and his/her terrain. They, however, continued to monitor each other's inventories of collected items and each other's retrieving attempts to guide their collecting actions. This mutual monitoring enabled them to describe the current problem state, for example, by calling out "no more screw (.) I have a screw" as a reaction to another's attempt at picking up one too many parts of this kind (which would have resulted in failing the level). In this way, they also continued displaying their orientation to the shared goal of accomplishing the task together.

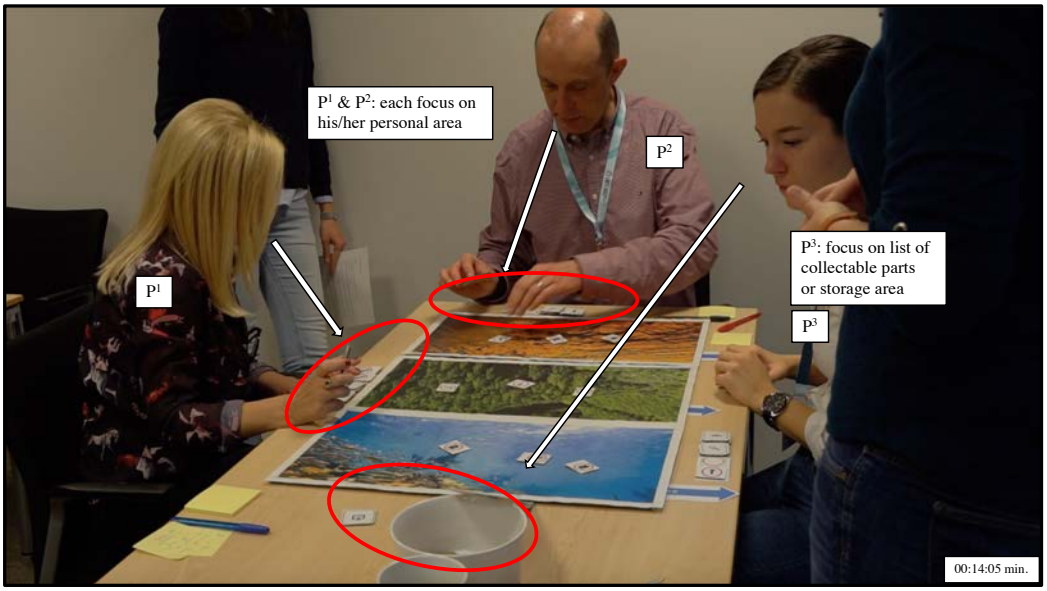


Figure 4. Divided visual focus of the participants

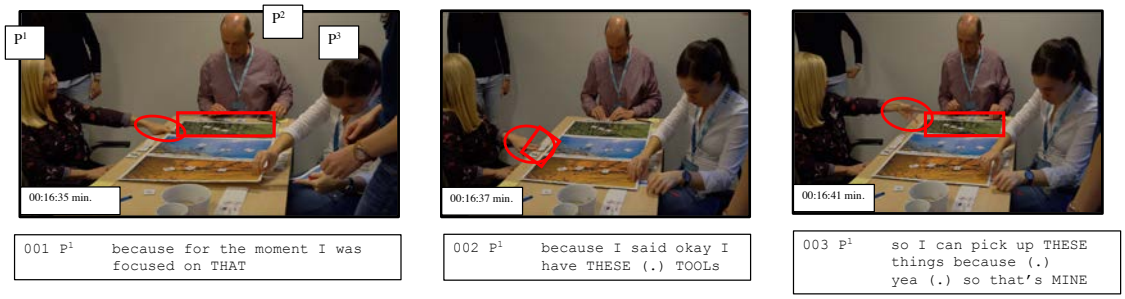


Figure 5. Making terrain-specific adherence accountable

Level 3 introduced the rule that an extra tool was needed to retrieve a part. Thus, most of the times, the competencies of two participants were needed to retrieve the requested parts. In response to this new constraint, participants coordinated among themselves to mobilize the appropriate tools (see figure 6). P^1, with the assistance of P^2, retrieves the information about the supplementary tool which is needed to collect a part from her terrain (lines 1-3), P^2 announces that he disposes of it (line 5) and the two participants jointly retrieve the part via a simultaneous mobilization of their respective avatars (lines 6-9).

```
001 P1    what what what (.) I need my glasses    ((P1 pulling part-card close to her face to decipher))
002       what's that                             ((P1 showing part to P2))
003 P2    showel(.) showel
          ((P2 reading out  word on part-card, joint group focus on part))
004 P1    okay (.) okay I have one                ((all checking personal tools in front if them))
005 P2    I can give you a showel
006 P1    so I (.) can (.) take
          ((P1 & P2 reaching @relevant part with both avatars, joint group focus on part and collection activity))
007 P2    okay
008 P1    I have one (.) I can take
009 P2    okay                                    ((P1 & P2 retraction from collection move))
```

Figure 6. Simultaneous conduct of simultaneous retrieving action

Throughout this level until the end, participants remain mutually engaged to successfully complete the level and so display their goal orientation; call out what tool or part is needed and advise caution, thus, pointing to the problem state; and show that they are aware of how to solve the ongoing problems, for example, by

announcing that they dispose of the needed tool or by suggesting to take other tools from the storage area.

Scenario 2: Growing crops

After a trial and error phase (in level 0) participants figure out together the appropriate order in which the farming resources have to be used to grow the crops successfully (tractor, seeds, water, fertilizer). The discovered procedure, which requires the sequential mobilization of the distributed competencies, becomes then available and recognized as a shared routinized problem-solving action to accomplish the tasks in all the levels. Overall participants establish a joint focus oriented at the field where the procedure is being applied, so that they can coordinate their actions to place the right farming resources in the right spot at the right time (see figure 7). At the end of the last level the procedure is further rationalized in the sense that a participant no longer waits until a field is finished but immediately moves on to the next one to apply his farming resource.

Figure 7. Joint focus

A perhaps more elaborate moment of JPS construction occurs, when the participants are challenged by new task constraints in level 1 (neighbor fields may not contain the same crop) and in level 2 (given amounts of different crops have to be harvested). On suggestion of one of the participants the seed bags are used during both levels as a planning aid to visualize distribution possibilities without actually initiating the procedure (figure 8). In this way potential solutions were shared, discussed, agreed upon and then implemented.

Figure 8. Mobilizing tangibles for visualization and planning purposes

So, again we could witness how mutually engaged participants displayed their shared commitment to accomplish the tasks, their collective awareness of the challenges of the tasks and of how to tackle them. As in the previous scenario the central space is fragmented, but there is no personalized appropriation of the fields by the participants. Probably, this is not solicited because the individual competencies are not tied to the fields.

Scenario 3: Collecting garbage in the sea

During level 0 a situation occurred demonstrating that designers have to give special consideration to the allocation and organization of participants' competencies. As outlined above, the operation of the steering directions was distributed in a complementary way, meaning that the participants had to coordinate among themselves to sequentially operate their respective directions to reach the previously negotiated destination. Due to the location of the harbor in the Southwest (levels 0 and 1), where the ship departed, and the location of most of the garbage items in the East, P^1 and P^3 controlled all the required movements to reach the related locations (figure 9). This combination of circumstances solicitated a close mutual engagement among P^1 and P^3 to select a destination and to collectively steer for it, but it also solicited a disengagement on the part of P^3, who made this explicitly accountable (figure 10).

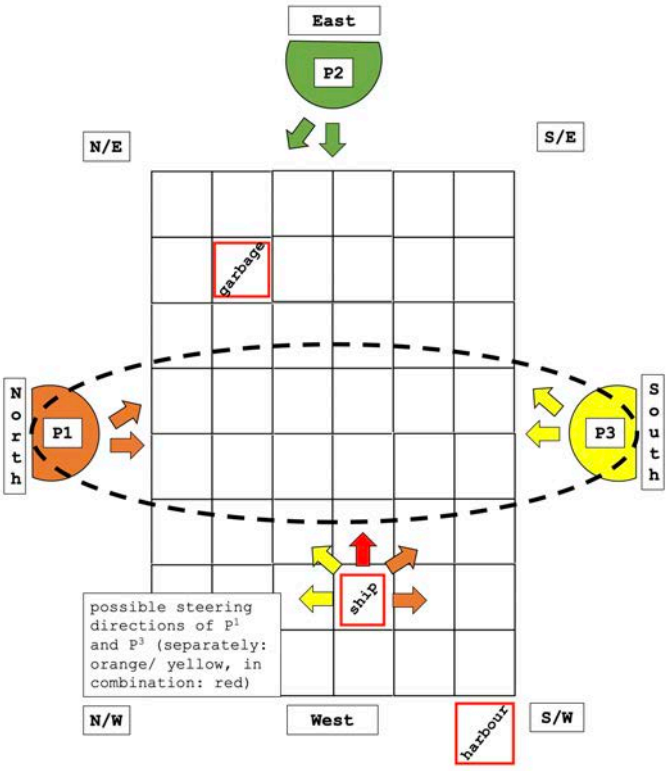


Figure 9. Distribution of steering directions leading to a temporary exclusion of P^3

Figure 10. Displaying being excluded

As long as there were no restrictions placed on the length of the route via fuel consumption (level 1 and 2), participants rather quickly agreed on the destination to target and moved the ship accordingly. During the steering they monitored one another's actions and sometimes prompted the participant, whose turn it was. During a spontaneous exchange between two levels, participants pointed out that they were instructing one another to do the requested steering, a conduct that was made possible through the general visual access to everyone's competency. This observation led to the concern that - at least in theory - one person alone could plan the trajectories and, all along, instruct the others accordingly. It is very unlikely that this organization of conduct would contribute to the establishment of a joint problem space.

Level 2 introduced a new rule (which was maintained for level 3), namely that every movement (in any direction) consumed 1 unit of fuel. 20 units were available in the ship's fuel tank (visualized through a gauge) and collectable refills (5 units) were located in the ocean. In response to these new constraints and in order to accomplish the task (shared goal), the participants mutually engaged in long planning and discussion phases where they considered various possible routes, carefully weighted them, and agreed upon a trajectory. Finally, they carried out the latter, while monitoring one another. During these phases, participants described the problem state, for example, by highlighting the current fuel limitations and the steering restrictions for the route under scrutiny (figure 11), and they displayed their awareness of the available problem-solving actions, for example, by pointing to an interesting target area containing a high concentration of collectable garbage and being in proximity of a fuel refill (figure 12); or by counting and verbalizing the steps to test a potential itinerary.

Figure 11. Problem-oriented sharing of individual steering options

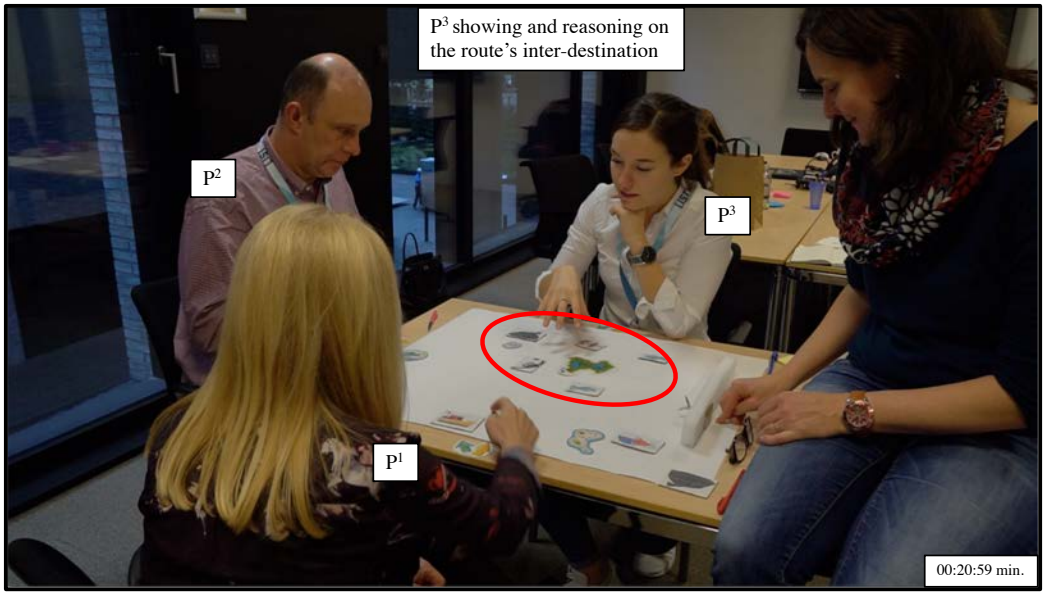


Figure 12. Highlighting an appropriate destination

After completing level 2, the participants displayed in an off-scenario discussion, that they were aware of these extensive and demanding planning phases. Indeed, contrary to the previous scenario, where the seed bags were spontaneously used as an organizing tool with regard to the crop-to-field allocation, the collecting garbage scenario did not provide an artefact that could be used to mediate/facilitate the decision-making process with regard to the best route to take.

Conclusion

The implementation and evaluation of paper prototypes in the context of a multidisciplinary design workshop was the first design step of an iterative research process, that aims at developing and investigating a TTI-mediated joint problem-solving activity (Sunnen et al., 2018). Although a paper prototyping cannot fully

simulate a computer interface, with regard to crucial features such as the provision of instant feedback and multitouch manipulations[7], the results we gained from our investigations will provide valuable insights to inform the choice of a scenario and the design of the upcoming digital prototypes.

Through our exploratory analysis we could show that collaborative conduct was elicited by all three scenarios. We could further highlight that the design of the task, the physical organization of the tabletop space and the distribution of complementary competencies have to be considered as intertwined design aspects, that are highly consequential on participants' collaborative conduct in TTI-mediated joint activities. Through the introduction of supplementary constraints, the tasks in the different scenarios became more challenging with regard to coordination and planning. With regard to the latter, it can be said that the requirements increased substantially from the first to the third scenario and solicited an appropriate and engaging joint response from the participants. In the first scenario ('damaged spaceship'), the additional constraint was implemented through a modification of the user competencies which rendered the participants' retrieving actions interdependent and synchronous. As we could observe, this entailed mutual monitoring and engagement. The second and the third scenario ('growing crops' and 'collecting garbage in the sea') required a sequential mobilization of the competencies, and solicitated coordination efforts and the establishment of a joint focus. The third scenario further teaches us that the complementary competencies have to be carefully thought through to elicit mutual engagement among all the participants in a balanced way. The organization of the tabletop space was particularly 'intriguing' in the first scenario, where the central space was threefold. This spatial arrangement, being bound to the competencies, did not facilitate the construction of a joint visual focus but did not impede the construction of a shared problem space either. The joint visual focus was restored when participants' competencies became interdependent.

A major design challenge is to expand the role of the TTI in the mediational framework of the joint problem space to bring forward the added value of the TTI. By that we mean that the TTI should become a powerful resource to be embedded in and interweaved with participants' joint meaning making processes. This aspect became most noticeable during the extended and demanding phases of the last scenario. Indeed, the tabletop did not provide the participants with facilitating means to keep and display uttered potential solution-oriented steps (for example, a hypothetical trajectory). Such a feature would make those contributions available for re-integration and transformation in the joint problem-space, thus, supporting a crossed backward-forward oriented joint decision-making process regarding the actions to take to achieve the shared goal.

---

[7] Participants sometimes made this jokingly accountable by saying "the computer is slow" or "it's a single processing computer".

Acknowledgments

We thank the Luxembourg National Research Fund (FNR) for funding this research under the CORE scheme.

References

Ioannou, A. and Antoniou, C. (2016): *Tabletops for peace: technology enhanced peacemaking in school contexts*, in International Forum of Educational Technology and Society.

Mercier, E. and Higgins, S. (2014): *Creating joint representations of collaborative problem solving with multi-touch technology*, in Journal of Computer Assisted Learning, vol. 30, no. 6, pp. 497–510.

Price, S. (2013): *Tangibles: Technologies and Interaction for Learning*, in Price, S., Jewitt, C. & Brown, B. (eds.), The SAGE Handbook of Digital Technology Research, pp. 307-326, London: Sage.

Fernaeus, Y. et al. (2008): *Towards a New Set of Ideals: Consequences of the PracticeTurn in Tangible Interactio*n, in Proceedings of the 2nd nternational Conference on Tangible and Embedded Interaction, pp. 223-230, New York.

Scott, S.D. et al. (2004): *Territoriality in Collaborative Tabletop Workspaces*, in Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work (CSCW '04), pp. 294-303, ACM, New York.

Speelpenning, T., Antle, A. N., Doering, T. and van den Hoven, E. (2011): *Exploring How Tangible Tools Enable Collaboration in a Multi-Touch Tabletop Game*, in LNCS, vol. 39, no. 3, pp. 605–621.

Fleck, R. et al. (2009): *Actions Speak Loudly with Words: Unpacking Collaboration around the Table*, in Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, pp. 189-196, ACM, New York.

Yuill, N. and Rogers, Y. (2012): *Mechanisms for collaboration: A design and Evaluation Framework for Multi-user Interfaces*, dl.acm.org.

Stanton, D. et al. (2001): *Classroom Collaboration in the Design of Tangible Interfaces for Storytelling*, dl.acm.org.

Woodward, J. et al. (2018): *Investigating Separation of Territories and Activity Roles in Children's Collaboration around Tabletops*, dl.acm.org.

Roschelle, J. and Teasley, S.D. (1995): *The Construction of Shared Knowledge in Collaborative Problem Solving*, in O'Malley, C.E. (ed), Computer-Supported Collaborative Learning, pp. 69-197, Berlin: Springer.

Sarmiento-Klapper, J. W. (2009): *The Sequential Co-Construction of the Joint Problem Space, Studying Virtual Math Teams*, pp. 83-98, Springer, Boston.

Sunnen, P., Arend, B. and Maquil, V. (2018): *ORBIT-Overcoming Breakdowns in Teams with Interactive Tabletops*.

Mercer, N. (1995): *The Guided Construction of Knowledge*, Clevedon: Multilingual Matters.

Vass, E. and Littleton, K. (2010): *Peer Collaboration and Learning in the Classroom*, in Littleton, K. et al. (eds.), International Handbook of Psychology in Education, pp. 105-135, Bingley: Emerald.

Cohen, E.G., and Lotan, R.A. (2014): *Designing Groupwork*, 3rd Edition, New York: Teachers College Press.

Piper, A. M., O'brien, E., Morris, M. R. and Winograd, T. (2006): *SIDES: A Cooperative Tabletop Computer Game for Social Skills Development*.