ISSN 2510-2591

Reports of the European Society for
Socially Embedded Technologies

volume 2 issue 1
2018

Proceedings of 16th European Conference on Computer-Supported Cooperative Work - Exploratory Papers

Guest Editors

Tom Gross, University of Bamberg (Germany)
Michael Muller, IBM Research Cambridge (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

Koch, M.; von Luck, K.; Schwarzer, J.; Draheim, S. (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, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2018_3

The Novelty Effect in Large Display Deployments – Experiences and Lessons-Learned for Evaluating Prototypes

Michael Koch[1], Kai von Luck[2], Jan Schwarzer[2], Susanne Draheim[2]
[1]Universität der Bundeswehr München
michael.koch@unibw.de
[2]Hamburg University of Applied Sciences
{jan.schwarzer, kai.vonluck, susanne.draheim}@haw-hamburg.de

Abstract. *This exploratory paper addresses the novelty effect in large display field deployments by combining findings from both the existing body of knowledge and our own research. We found that the novelty effect is prevalently present on two occasions: (a) immediately after a new system is deployed in a new environment, and (b) in reoccurring situations, when changes are made to an existing system. Both instances share similarities such as a system's higher usage during a particular time frame. However, we also observed that their individual reasons to occur are multifaceted. The present work's main contribution is twofold. Firstly, the paper outlines related literature regarding the novelty effect, particularly in CSCW and HCI. Secondly, the paper illustrates the effect's complex nature and suggests explicit means that should be considered in related research endeavors.*

Introduction

Understanding the practical application of technology is an important part of CSCW and HCI research. Lab-based studies, however, provide little information on how technologies can be adopted and applied in real-word contexts, or if they

are ever likely to be adopted at all. Field deployments or pilot implementations which evaluate the impact of new technologies in real-world environments are needed to understand how people utilize technology in their everyday lives (Hertzum, Bansler, Havn, & Simonsen, 2012; Siek, Hayes, Newman, & Tang, 2014). Specifically, field deployments and pilot implementations provide rich data about how closely a concept meets the target demographic's needs and how users accept, adopt, and appropriate a system in actual use over time. Field deployments can also be used to validate a concept or prototype – both for systems based on well-established, recognized needs, and for groundbreaking technological innovations.

Overcoming the novelty effect is a major challenge in determining a new technology's practical application potential. For example, research shows that it can take up to six months for a new behavior, such as the full adoption of a new technology, to become habit (Prochaska & DiClemente, 1982). In that time, social pressures or trends that drive initial use may wear off over time, or unforeseen issues with, e.g., scalability or maintenance may arise.

When planning research, one must address the question of how long an innovation should be deployed and evaluated in order to avoid skewing research outcomes. In our research on semi-public and public displays in collaborative environments, we encountered this issue quite often, and began searching for appropriate answers. We found some generic guidelines for evaluating public displays, such as (Alt, Schneegaß, Schmidt, Müller, & Memarovic, 2012), but little practical guidance regarding novelty effect's impact in various deployment contexts, or how long one should evaluate a new system in order to mitigate the effect's influence.

This paper sets out to explain the patterns we observed in our research and to derive some lessons-learned in order to assist researchers in similar studies. The following is structured as follows: firstly, mentions and findings about the novelty effect are collected from multiple disciplines, including CSCW and HCI; secondly, findings regarding the novelty effect in our studies and experiments with semi-public and public displays are presented; thirdly, these findings are discussed, and a set of preliminary dimensions for summarizing influencing factors on the novelty effect is described; finally, we conclude by providing direction for future research.

The novelty effect in different disciplines

A formal and comprehensive definition of the novelty effect is currently missing. However, Wikipedia provides a summary of what appears to be a commonly-accepted definition:

> "The novelty effect, in the context of human performance, is the tendency for performance to initially improve when new technology is instituted, not because of any actual improvement in

> learning or achievement, but in response to increased interest in the new technology." ("Novelty Effect", 2017)

Another definition for the novelty effect is provided in (Ott, 2018):

> "The curiosity effect is the decreasing intensity of use of a new or recently updated technical component in a sociotechnical system, which is due to the increasing integration in the workday life and the loss of interest by the social actors that is accompanied by this." (translation by the authors)

To summarize, the novelty effect is an increased motivation to use something, or an increase in the perceived usability of something, on account of its newness. When novelty eventually fades, usage patterns and/or perceived usability changes.

Novelty effect and Hawthorne effect

Psychologists noticed a phenomenon similar to the novelty effect in the 1930s, when several changes affecting working conditions in the Hawthorne Works factory were implemented and evaluated for their effects on productivity. The studies found that it did not matter what had changed – any change in the workplace produced a temporary boost in productivity (Landsberger, 1958). The Hawthorne effect states that those who perceive themselves as members of an experimental or otherwise favored group tend to outperform control groups, even in the absence of applied variables. Thus, the Hawthorne effect describes the effects of knowingly being observed. In contrast, the novelty effect stems solely from an innovation's newness with no regard to observation or a user's knowledge thereof.

Novelty-Encoding Hypothesis in medical psychology

In medical psychology and neuroscience, there is a so-called "Novelty-Encoding Hypothesis" proposed by (Tulving & Kroll, 1995), suggesting that the encoding of online information into long-term memory is influenced by its novelty, and that novelty increases recognition performance. This hypothesis was confirmed in several different settings, for example in (Kormi-Nouri, Nilsson, & Ohta, 2005) and (Poppenk, Köhler, & Moscovitch, 2010). This effect is closely related to what we have described as the novelty effect, but again, it is not the same. The hypothesis focuses on recognition performance of memorized information, while the novelty effect focuses on usage patterns as a direct result of an object's initial installation or changes to its state (e.g., adding new features).

Novelty effect in educational research

First quantifications of the novelty effect can be found in educational research. There are several studies concluding that students learn better when working with computer-based learning material (Kulik, 1994). However, (Clark & Sugrue, 1988) already noted that achievement gains can be attributed to an increase in students'

attention, but diminish once students become familiarized with the new medium. They found that novelty effects caused an increase in standard deviation averages of 30% concerning achievement gains, but then decays to a smaller margin after eight weeks. (Krendl & Broihier, 1992) presented a study which showed that improvement is even lower, or absent altogether, when studies are conducted over longer periods of time. Their study covered a period of three years, and the findings clearly demonstrate evidence of novelty effect. Students' preference for computers, as well as their perceptions of learning from technology, declined significantly during the three years. Contrary to expectations, the perceived difficulty of using computers remained stable.

Novelty effect and technology acceptance

An approach to generalize the results from educational research would be to examine the relationship between a utility's novelty and its usage patterns. Initial thoughts on this topic can be found in an online article by Clive Thompson (Thompson, 2014). He elaborates on "why a new high-tech tool makes you suddenly more productive or creative – until it doesn't".

Enterprise Social Networks (ESN) are one particular category of online collaboration tools in which the novelty effect has been documented. For example, (Glaser & Ebersbach, 2013) attribute changes in wiki usage patterns to the wiki's fading novelty and the subsequent dwindling of user curiosity. The tool "becomes part of the gray ordinary working day and loses attractivity".

While some reports indicate steady growth in ESN usage, albeit at widely differing rates, (Koch & Bentele, 2011) state that approximately 58% of potential users visit the ESN daily. Others present a steady decrease in usage until an incentivized challenge was released, a reminder was sent, or new features were introduced (e.g., (Müller & Stocker, 2012)).

When investigating models explaining the influence of novelty on adoption, studies on the determinants of acceptance provide preliminary insights. The Technology Acceptance Model (TAM) is one of the most widely used theoretical frameworks that addresses user acceptance or rejection of a new technology (Davis, Bagozzi, & Warshaw, 1989). Novelty can influence the factors of *Perceived Usefulness* and *Perceived Ease of Use* in TAM. It might also influence the factor of *Perceived Enjoyment* (Merikivi, Nguyen, & Tuunainen, 2016). For example, (Webster & Ho, 1997) tested the effects of increasing variety by incorporating new sound effects and animations into a presentation. (M.-H. Huang, 2003) employed the term "novelty" to refer to aspects of a system that users perceived as surprising or unfamiliar. Additionally, a tool's novelty can be seen as a user experience dimension, e.g., to be measured using the User Experience Questionnaire (UEQ) (Laugwitz, Held, & Schrepp, 2008).

The novelty effect in CSCW and HCI research

Several studies exist in CSCW and HCI research that mention the novelty effect in some way. For example, Huang et al. discussed various factors for success and failure with large-display groupware systems (E. M. Huang, Mynatt, Russel, & Sue, 2006). They found that several deployments were characterized by strong initial usage followed by decreasing, more sporadic use.

In their methodology section, Gallacher et al. argued that their artifact was deployed for four weeks in order "to provide enough time to investigate the initial novelty effect" (Gallacher et al., 2015). In the quantitative analysis, they reported that the initial spike in usage declined to a stable level by the second week.

Guerrero et al. suggested that "the novelty factor" likely impacts the motivation to use their artifacts (Guerrero, Ayala, Mateu, Casades, & Alamán, 2016). They concluded that they would need to evaluate their solution over a longer duration in order to determine the extent of this effect.

Hosio et al. stated that many of their displays were deployed in the same place for several years and thereby outlasted their perceptions as novelties, "which is important, as inserting novel technology in public often leads to strong novelty effects and bias in the actual usage." (Hosio, Goncalves, Kostakos, & Riekki, 2014)

As Hazlewood, Stolterman & Connelly noted, "The four-week mark was selected because it was predicted that this was adequate time for most people to have noticed the display, and for the initial novelty factor to wear off." (Hazlewood, Stolterman, & Connelly, 2011)

One example wherein the novelty effect received attention is a study on Chained Displays (ten Koppel, 2011; ten Koppel et al., 2012). The authors investigated to what extent a novelty factor was present following the installation of a game designed to study varying configurations of large interactive screens. To address the novelty effect, they included a post-game questionnaire asking users if they had had seen the installation already, and if it was their first time playing the game. The results (day 1: 93.8%, day 2: 73.1%, day 3: 59.1%) indicated that general awareness of the installation progressed throughout the study; however, the question regarding user's first gameplay experience showed that 72.7% of users who played the game were still new users on the third day.

Additional implicit information on the novelty effect can be found in reports on lengthier studies in the context of playable cities. One example is the UBI Hotspots in Oulu (Ojala et al., 2010, 2012). The authors stated in their 2012 review:

> "We have also observed the novelty factor when we introduce new system features or release a major system upgrade. In both cases, use spikes but then gradually decreases. The effects of novelty vary across instances, but we have consistently observed its impact to some degree [...]"

Finally, the novelty effect can be linked to the idea of "display blindness" discussed and measured in large screen deployments, e.g., "[...] numbers early in the deployment allow a potential novelty effect to be quantified. Numbers at the end of the deployment are most interesting, since they reflect the degree to which

a deployment manages to overcome display blindness beyond the novelty effect." (Memarovic, Clinch, & Alt, 2015)

Discovering the novelty effect in large (semi-)public display deployments

Above, we presented some examples of public display research where the novelty effect was either explicitly mentioned or rudimentarily measured. Motivated by the questions surrounding the novelty effect, we closely examined the data from our own field deployments and then identified metrics that would help us investigate this effect more thoroughly. In the following examples, we present selected results and insights gained from these studies.

(Semi-)public information displays in the CommunityMirror project

In the CommunityMirror project, public displays present non-work-related information deemed nonetheless interesting or useful in the workplace. The screens are installed in semi-public places, e.g., beside elevators, in break rooms, and other social areas where people congregate. By displaying such information, these interfaces can help increase visibility of happenings within the organization (awareness) and facilitate the "accidental" discovery of relevant information without employees having to look it up explicitly (serendipity) (Ott & Koch, 2012).

Experiment 1: IdeaMirror (Blohm et al., 2010)

In this study, we deployed a large interactive screen (IdeaMirror) near elevators and in a business incubator's cafeteria for six weeks. A set of customer-generated ideas was presented to 198 employees from 59 start-ups for voting and commenting purposes.

The actual usage of the IdeaMirror was evaluated by analyzing log file data. In the first days of testing, initial interest in interacting with the new technology was observably strong. Interest decreased over time but spiked following external stimuli (e.g., mentions in an email newsletter), followed again by eventual decrease. Thus, a wave pattern is present in the usage data (see Figure 1).

Experiment 2: Focus on getting attention (Lösch, Alt, & Koch, 2017)

In this study, we deployed an interactive installation consisting of (1) a direct-touch information display with which users could interact using touch gestures, as well as (2) larger projected screens, referred to here as *hallway displays*. The interactive installation was built alongside a pathway. The hallway displays were dynamic, i.e., arrangeable in various configurations, and allowed the evaluation of multiple deployment setups within the interactive hallway.

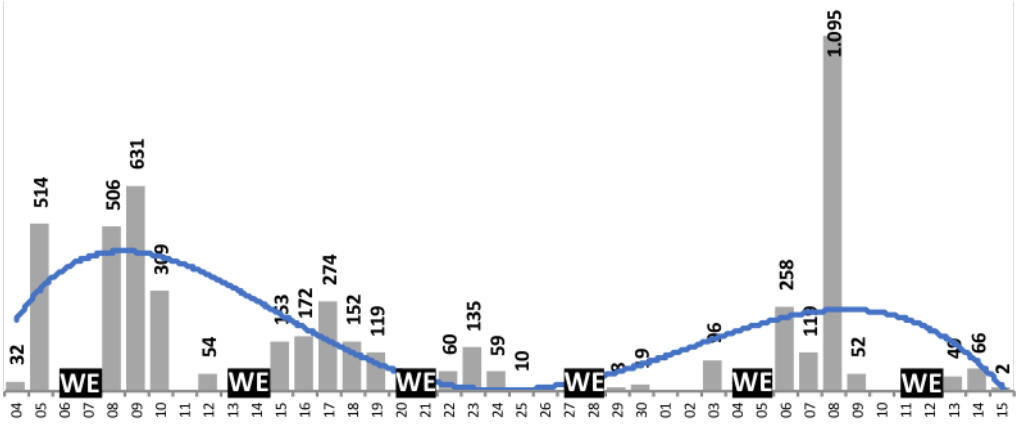

Figure 1. Timely distribution of interactions with the IdeaMirror during the whole field test.

The interactive installation was deployed in a university canteen during lunchtime for six weekdays over two consecutive weeks (from Tuesday to Thursday, respectively). In the field study, five different variants of the interactive hallway (v1-v5) were evaluated. During the six days of our deployment, we changed the setup daily.

To better understand the novelty effect in our setup, we decided to utilize one configuration (v5) of the hallway displays twice – once at the beginning and once at the end of the deployment. Note that the canteen is the main location on campus to have lunch; hence, it can be assumed that many students and staff members visited daily, and crossed the long, one-sided hallway during the initial deployment. The result: the percentage of passers-by that stopped at the installation decreased from 24.7% on the first day to 14.2% on the last day. The interaction rate decreased from 17.1% to 9.9%. The average duration of interaction was also higher (31 s) on the first day than the last (23 s).

Experiment 3: INFMirror

In this study, we deployed a large interactive screen in the semi-public area of a university building next to info-boards that many students and staff members walk by daily.

The screen showed information about people in the department, publications and projects by department members, as well as upcoming departmental events. The display's standard view only visualized items with titles and images. By touching an item, users could access additional information.

We measured direct interaction with the screen over twelve weeks. In the data, there was an observable peak in the first week (70 interactions/week) – followed by lower rates of interaction in subsequent weeks (about 30 interactions/week). This could be explained by the novelty effect – people interacted with the screen out of sheer curiosity rather than having an actual interest in the displayed content.

Ambient Surfaces: a long-term study of semi-public displays for co-located agile software development teams

In recent years, we collected data using experimental interactive display installations in real-world environments, such as in (Schwarzer & von Luck, 2012). Different scenarios and runtimes were applied depending on the deployment's application context and intent; nonetheless, it was consistently evident that the novelty effect influenced the intensity of use in the early stages of each study. Our ongoing research has produced similar results, and so we began to thoroughly document parameters that we believed were influencing this phenomenon (e.g., the release dates of new features and updates).

The Ambient Surfaces project aims to provide the project partner's agile software development department with large interactive displays (see Figure 2). In collaboration with the department, these systems were revised over the years to maintain their value, e.g., by continuously integrating new or upgrading existing information views such as GoCD[1]. The intent is to (a) access and display relevant information from the company's intranet and (b) display this information in a high-traffic common room on large screens (≥ 46 inch), thereby encouraging informal gatherings and discussions. The displays contain multiple information layers when accessing their contents. Some views display basic information such as a website; in these cases, one layer is sufficient to present the information. In other instances (e.g., GoCD), the volume of data required us to build several view layers and incorporate user interface elements such as scrollable lists. In addition, different colors and animations were utilized to attract attention (e.g., when automated builds failed to execute). The first Ambient Surface was installed in February 2014 followed by the installation of a second system in August 2015. This was due primarily to usability issues resulting from the amount of relevant information accessible from the various development tools in the department.

A mixed-design grounded theory is used in our research. Data collection techniques incorporate qualitative as well as quantitative data, including touch interactions, observations, interviews, team calendar information, field notes, email feedback, and visual data recorded by two Microsoft Kinect cameras. Initial results were presented in (Schwarzer et al., 2016). By contrasting various data parameters, we set out to understand how, when and in what contexts the Ambient Surfaces are being utilized, what information is relevant, and how this correlates to the department's "formal frame" of communication (e.g., daily stand-up meetings). Throughout the study, by analyzing touch interaction logs, two phenomena were dominant in the material with respect to the novelty effect: (a) an increase in interactions following the systems' initial deployments, and (b) reoccurring spikes in interaction during or immediately following system updates.

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

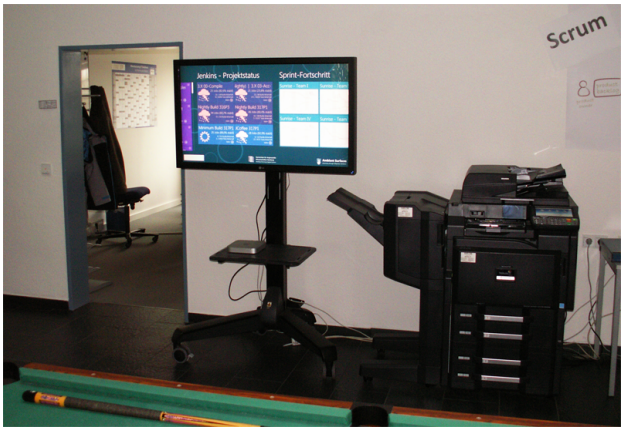

Figure 2. Ambient Surface setup in a common room area (March 2015).

In 2014, the mean touch-down events[2] per week first fell below the all-year mean (485.22) in Week 18, which was the 11th week of the field study (see Figure 3). Out of 21,835 touch-down events in 2014, weeks 8–17 accounted for approximately 73% (15,949) of them. It is worth noting that, due to some initial technical issues, the system went operational on Friday during Week 8, which explains why the largest number of interactions ever recorded occurred in Week 9. Furthermore, as a result of third-party component compatibility issues, it was not possible to deploy a web browser component for displaying content from a wiki website until Week 11. In addition, out of roughly 75 employees who worked in the same building, only a few were familiar with the Ambient Surfaces prior to deployment. A small team of representatives (including management personnel, Scrum Masters and developers) were involved in the final phases prior to February 2014. The use of such physical artifacts was also novel in the department. Another factor worth considering stems from the circumstance that not all employees were always present. While discussing touch interaction patterns in a group interview, it was apparent that there were several reasons for absences in the workplace, such as illnesses, trainings, and holiday seasons. With respect to the all-year mean interactions, this may explain the lengthier and greater amount of interactions during the first ten weeks when compared to examples from related literature (Hazlewood et al., 2011). However, the first four weeks showed the highest quantity of interactions.

[2] In the following, only touch-down events occurring between 7.00 a.m. and 8.00 p.m. were considered.

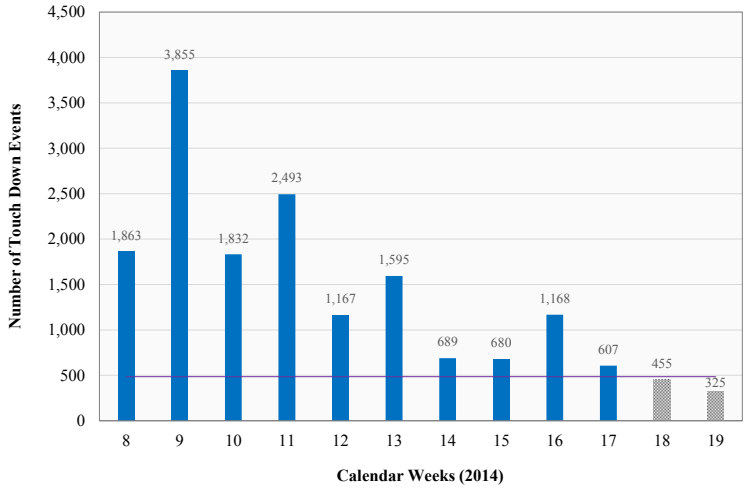

Figure 3. Touch down events per week (8–19) in 2014 (all-year mean=485.22).

Additional patterns were observable before, during, or after changes were applied to the systems (e.g., replacing or updating an information view). The following exemplarily displays three events which occurred during the first 20 weeks of 2015 (see Figure 4).

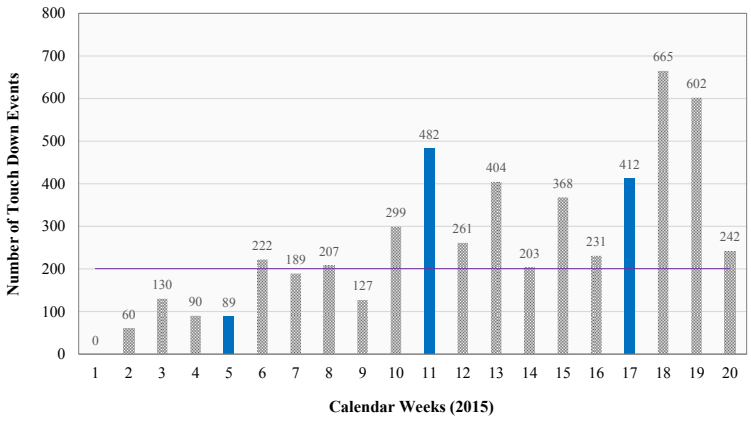

Figure 4. Touch-down events per week in the first 20 weeks of 2015 (all-year mean=200.66).

The first update was on account of the department's use of Jenkins[3] for continuous integration tasks. Occasionally, new product versions are created and new projects are generated in Jenkins therefor. These projects must be registered in a configuration file to be displayed on the Ambient Surfaces (as indicated in Week 5). As a result, the number of interactions per week increased throughout the following weeks, possibly due to the increased usefulness of the presented contents as development progressed over time. These change requests are predominantly communicated by Scrum Masters.

[3] https://jenkins.io/

The second update was the result of direct discussions with developers, and a new information view was added in Week 11. The view displayed a table with known bugs and corresponding teams who were responsible for resolving the issues. By Week 13, this view was revised in collaboration with developers. The modification also resulted in a higher number of interactions, possibly attributable to an increase in content quality. It should be noted that the collaboration and revision process with the developers likely affected the total number of interactions. For example, developers reviewed recently deployed user interface updates with colleagues and interacted with the surface.

The third update followed observations, informal discussions, and a subsequent group interview; the whole system then underwent considerable revision in Week 17. Some information views were removed, a new view was added, and another was updated. Three Jira[4] RSS feed visualizations were removed due to the high volume of dynamic information displayed throughout the day. Respondents found it difficult to follow; thus, the Jira visualizations were of little value. Users preferred to utilize the display's available space more efficiently by displaying other helpful contents, which included another Jenkins view to better distinguish between product versions and their corresponding Jenkins projects. Furthermore, a Jira-based burndown chart was added to display a product version's progress. All of these changes resulted in an increase in interactions. It is worth noting that, prior to the revised deployment on Sunday of Week 17, the system was evaluated by at least some members of the management team (who sent us an email with final instructions on Friday), which may explain the increased interactions in Week 17 and suggest an increase in the perceived value of the content displayed in subsequent weeks.

We consistently observed patterns similar to the examples above when deploying other updates and revisions to the Ambient Surfaces throughout the entire study.

Lessons-learned for designing prototype evaluations

As noted in the introduction, the novelty effect must be considered in any research dedicated to understanding the potential use of novel technologies. In the following section, we distilled some lessons-learned and crafted recommendations for future researchers. We found that the novelty effect was predominantly present in two scenarios: (a) initially, after a system's deployment, and (b) when changes are made to the system's state. The results suggest that formulating conclusions with respect to the novelty effect can present more challenges than researchers may initially anticipate.

---

[4] https://de.atlassian.com/software/jira

How long does the initial novelty effect last?

Existing research recommends conducting deployments for a sufficient period of time in order to ensure that results are not tainted by the novelty effect. For example, (Hazlewood et al., 2011) emphasized "*Run[ning] for an extended time"* to ensure that the display was adopted and fully integrated into participants' daily routines. They reported that the timeframe adequate for most people to notice the display was four weeks, i.e., the novelty effect diminished during that time. Additional evidence proposes that it can take up to six months to mitigate the novelty effects of new technology (Prochaska & DiClemente, 1982). We found that the novelty effect took ten weeks to subside in our Ambient Surface project. One week, however, was sufficient for the usage patterns to stabilize in Experiment 2 of the CommunityMirrors project.

Further examination of factors influencing the novelty effect is necessary to provide guidance for researchers. For example, the CommunityMirrors and the Ambient Surfaces can be categorized as voluntary and mandatory applications, respectively. This distinction was relevant in the context of enterprise social networks, where maintainers commented that they did not see decreased usage because daily log-ins were mandatory. However, this approach could have its own limitations because, even with mandatory usage, users may demonstrate varying levels of engagement. More precisely, the novelty effect may be related to how deeply integrated a tool or technology is into an organization's operations. If this is the case, then the effect would be highly dependent on a tool's operational necessity, ranging from mission-critical (e.g., ERP systems) to occasionally useful (e.g., coworker status updates). The CommunityMirrors, for example, are not intended to be integrated into business operations. Finding ways to attract an individual's attention is already an inherent design challenge, and so the novelty effect quickly faded (e.g., Experiment 2). By contrast, the Ambient Surfaces are designed for integration into business operations by displaying only relevant, work-related information. Employees are more motivated to use them, thereby influencing the observable duration of the novelty effect.

How long does the reoccurring novelty effect last?

We found that the novelty effect is not only important at the beginning of a deployment, but also throughout its operational lifespan. Especially in long-term deployment studies, novelty is repeatedly caused by changing a system's state, e.g., adding new features, drawing attention to the installation and/or its existing features, or simply updating its content. For systems not designed for workplace integration (e.g., the CommunityMirrors), this reoccurring novelty effect, i.e., the increased attention and the subsequent usage spikes following content updates, is critical. The effect helps maintain user interest and bolster perceived usefulness.

As with the initial novelty effect, we were interested in learning more about its counterpart. In the Ambient Surfaces deployment, we found that usage pattern anomalies (i.e., compared to all-year mean analyses) were regularly observable because the systems needed frequent updating in order to provide continuous value for employees. Precisely identifying the root causes of usage spikes, however, is not always possible without the constant presence of a researcher in the workplace (e.g., when a new employee first encounters a system). In other cases, causes were easily identifiable, e.g., after adding new features developed in collaboration with select employees. Thorough observation of the effect can also help determine a feature's practical value. In one instance, we found that our *meeting reminder* feature was of little value – it was seldom used even in the novelty phase, and usage then rapidly dropped to zero.

Our data revealed varying magnitudes with respect to the reoccurring novelty effect. Some changes entail a higher usage in subsequent weeks, such as in Week 5 (see Figure 4); other changes suggest that this implication may not always be the case, such as in Week 11. We conclude that, even in reoccurring cases, the novelty effect can last several weeks, but additional research is necessary to substantiate these findings.

Summary and preliminary set of dimensions

Identifying the exact reasons for the novelty effect in both reoccurring and initial contexts is a highly complex procedure, as is determining when the effect begins to abate. It implies the need for resource-intensive data collection techniques such as in-situ observation. Furthermore, some influential factors are not feasibly measurable, such as parental-leaves and holiday seasons.

In our research, the novelty effect is relevant for both deployments, but may not necessarily be relevant for related studies. For example, initial identification of the effect helped to validate the Ambient Surfaces' perceived value to employees for months and even years after the initial deployment. It also prevented us from presenting arbitrary conclusions. The reoccurring cases assisted, e.g., in determining a feature's long-term value. Table 1 summarizes our projects' findings in relation to the novelty effect.

Novelty Effect Occasions	Purpose	Duration
CommunityMirror project		
Initial	*E.g., determining the system's usefulness.*	*1–2 weeks.*
Reoccurring	*E.g., importance of constantly introducing new information for maintaining usage.*	*Still to be analyzed.*

Ambient Surfaces project		
Initial	*E.g., determining the systems adoption.*	*10 weeks (based on mean analyses).*
Reoccurring	*E.g., value of new features.*	*In some cases, several weeks. Further research necessary.*

Table 1. Novelty effect occasions, purposes, and durations in our research projects.

The duration of novelty-based behavioral change depends on a variety of contextual variables. Table 2 proposes a set of preliminary dimensions to classify deployment situations in order to assist researchers with classification of their own research endeavors and help identify factors that influence the novelty effect in their studies. In summary, we see two categories of dimensions combining (1) factors related to the user, usage, and value to the user and (2) factors related to novelty-inducing changes, including, e.g., changes to the setup, functionality, and the displays' contents in particular.

Regarding the first category, we found that an individual's *intrinsic motivation* (e.g., Scrum Masters' and management's interest in keeping the contents up-to-date) must be considered. Furthermore, the *number of potential users* (e.g., in cafeterias versus offices), the typical *intensity of use* (e.g., the number of touch inputs necessary for system interaction), the *required competence* (i.e., necessary training), and the *relevance for practice* (e.g., CommunityMirrors in contrast to Ambient Surfaces) have been identified as key influential factors.

In the second category we observed that the *frequency of change*, the *magnitude of change*, the *contents' presentation*, the *means for attracting attention*, and the *participatory development* process during a system's deployment or update are important considerations. As noted above, the Ambient Surfaces deployment utilizes various designs and interface components for displaying content, e.g., color schemes, animations, scrolling lists, and nested layers. The interfaces were regularly updated with new components and features to varying extents, sometimes developed in direct collaboration with employees. The implicit effects of applied interface components must be considered when investigating the novelty effect. Navigating through multiple information layers or scrolling through content, for example, will increase the number of logged interaction events. In one case, we received feedback that blinking red Jenkins items (indicating unexpected build failures) often raised the attention and curiosity of employees as to their cause. This was sometimes averted by the automatic update feature which fetched the latest data from the Jenkins server every ten minutes. Spikes in content navigation and scrolling events were nonetheless evident to a degree in our interaction log analyses. Furthermore, collaborative development has likely affected the degree of perceived novelty present in our research. For example, when the number of employees involved in collaborative development increases, we hypothesize that

the observable post-update novelty effect decreases, i.e., the number of interactions does not notably increase.

Influencing factors	Range of Value	
	From	*To*
(1) Influencing factors related to users		
Action orientation / Relevance for practices	*None (e.g., advertising)*	*Strong (e.g., dashboards)*
Intrinsic motivation for use	*Low*	*High*
Number of users	*Low*	*High*
Intensity of use	*Occasionally (e.g., ticket machine)*	*Regularly (e.g., dashboards)*
Required competence	*None (e.g., naive users)*	*Training (e.g., business users)*
(2) Influencing factors related to change (causing novelty)		
Frequency of change	*Occasionally*	*Regularly*
Magnitude of change	*Low*	*High*
Presentation (of content)	*Simple*	*Complex*
Means for attracting attention	*Few*	*Many*
Participatory development	*Without end users*	*With end users*

Table 2. Preliminary set of dimensions of influences and their attributes.

The CommunityMirror example can be categorized as follows: occasional use, naive users, weak action orientation, low intrinsic motivation, and a high number of users. The research questions focused primarily on drawing user attention and awareness. For the system's use case, the reoccurring novelty effect is strategically useful for achieving the intended goal, but the initial novelty effect can lead to misinterpretation of the data. The Ambient Surface deployment, however, can be categorized as follows: regular use, trained users, strong action orientation, high intrinsic motivation, and a high number of users. In addition, changes – including cases with considerable revisions (high magnitude) – were regularly deployed. The content depth ranged from simple to complex, and in addition, various means to draw attention and increase the value of displayed content were conceived and deployed, occasionally in direct collaboration with employees.

The novelty effect in both contexts (initial and reoccurring) is a relevant influencing factor when analyzing interaction data. Depending on the research questions, these patterns should be identified in the corresponding log data, and their affects accounted for to ensure unbiased analysis.

Conclusions

This paper presented and discussed the current body of knowledge regarding the novelty effect, particularly in the CSCW and HCI disciplines. Two denotations were used to depict the novelty effect on different occasions: (a) initially, when a system is deployed and (b) reoccurring, when a system is changed or updated. Examples were provided and possible impact factors were explained for both cases. Although this exploratory paper presents preliminary results, we believe our findings contribute to the discussion of the novelty effect in CSCW and HCI research by explicitly identifying observed instances of occurrence. Researchers, specifically those conducting studies on a system's adoption, may find valuable insights in our discussion, and practitioners can learn more from multifaceted questions regarding the introduction of new technologies in practical contexts.

Our research is not without limitations, which primarily stem from a conceptional and methodological level. CSCW and HCI research has only recently begun to investigate what *novelty* inherently implies. Novelty is a complex concept, highly dependent on intertwining variables such as application context and target demographics. Thus, promising directions for future research include exploring means to more precisely conceive the term novelty. One approach is to clarify the concept itself by distinguishing between characteristics such as new content, new interactive features, and new hardware. Another approach is to address the practical influence of novelty, and provide researchers with a standard for identifying and measuring the novelty effect's presence, and a means to account for the effect's influence on their own studies. We learned that qualitative tenets (e.g., thick description) helped in investigating the novelty effect, but research still fails to present sound methodological recommendations for evaluating prototype technologies in this regard. The intent of this explorative paper is to provide initial ideas for accomplishing this goal. We were not aiming at creating a model for the various factors that influence novelty and how these factors affect (initial) usage. However, technology acceptance research suggests that initial use involves alignment of artifact, work practices and organization. This results in the identification of a range of factors influencing use as defined in the TAM model (Davis et al., 1989). According to this model, a decline in usage may stem from decreased perceived usefulness and difficulty of use. We invite further conceptual discussion including technology acceptance models (e.g., TAM, UTAUT, or (DeLone & Mclean, 1992)) to clarify the novelty effect's plausibility.

Lastly, we call for more thorough research into the distinctions between related effects such as the Hawthorne effect. Such research is valuable for short-term experiments in real-world environments as well as the novelty effect's conceptualization.

References

Alt, F., Schneegaß, S., Schmidt, A., Müller, J., & Memarovic, N. (2012). How to evaluate public displays. In *Proceedings of the 2012 International Symposium on Pervasive Displays (PerDis 2012)* (pp. 1–6). New York, New York, USA: ACM Press. http://doi.org/10.1145/2307798.2307815

Blohm, I., Ott, F., Huber, M., Rieger, M., Glatz, F., Koch, M., … Krcmar, H. (2010). Extending Open Innovation Platforms Into the Real World – Using Large Displays in Public Spaces. In EURAM (Ed.), *Proceedings of the 10th European Academy of Management (EURAM)* (Vol. 2010). Rom: EURAM: European Academy of Management.

Clark, R. E., & Sugrue, B. M. (1988). Research on instructional media, 1978-1988. In D. Ely (Ed.), *Educational Media and Technology Yearbook 14* (pp. 19–36). Littletown, CO: Libraries Unlimited.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, *35*(8), 982–1003. http://doi.org/10.1287/mnsc.35.8.982

DeLone, W., & Mclean, E. R. (1992). Information Systems Success: The Quest for the Dependent Variable. *Information Systems Research*, *3*(1).

Gallacher, S., O'Connor, J., Bird, J., Rogers, Y., Capra, L., Harrison, D., & Marshall, P. (2015). Mood Squeezer: Lightening up the Workplace through Playful and Lightweight Interactions. In *CSCW'15 Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (pp. 891–902). ACM Press. http://doi.org/10.1145/2675133.2675170

Glaser, M., & Ebersbach, A. (2013). Reife Firmenwikis - Herausforderungen im späteren Prozess- und Wissensmanagement. Retrieved December 14, 2017, from https://blog.hallowelt.com/2013/02/05/reife-firmenwikis-herausforderungen-im-spateren-prozess-und-wissensmanagement/

Guerrero, G., Ayala, A., Mateu, J., Casades, L., & Alamán, X. (2016). Integrating Virtual Worlds with Tangible User Interfaces for Teaching Mathematics: A Pilot Study. *Sensors*, *16*(11), 1775. http://doi.org/10.3390/s16111775

Hazlewood, W. R., Stolterman, E., & Connelly, K. (2011). Issues in evaluating ambient displays in the wild: two case studies. In *CHI '11 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 877–886). ACM Press. http://doi.org/10.1145/1978942.1979071

Hertzum, M., Bansler, J. P., Havn, E. C., & Simonsen, J. (2012). Pilot implementation: Learning from field tests in IS development. *Communications of the Association for Information Systems*, *30*(1), 313–328. Retrieved from http://akira.ruc.dk/~mhz/Research/Publ/CAIS2012_preprint.pdf

Hosio, S., Goncalves, J., Kostakos, V., & Riekki, J. (2014). Exploring Civic Engagement on Public Displays. In *User-Centric Technology Design for Nonprofit and Civic Engagements* (pp. 91–111). Cham: Springer. http://doi.org/10.1007/978-3-319-05963-1_7

Huang, E. M., Mynatt, E. D., Russel, D. M., & Sue, A. E. (2006). Secrets to success and fatal flaws: the design of large-display groupware. *IEEE Computer Graphics and Applications*, *26*(1), 37–45.

Huang, M.-H. (2003). Designing website attributes to induce experiential encounters. *Computers in Human Behavior*, *19*(4), 425–442.

Koch, M., & Bentele, M. (2011). Rheinmetall: Unified Communication and Collaboration. In P. Schubert & M. Koch (Eds.), *Wettbewerbsfaktor Business Software* (pp. 229–246). München: Hanser.

Kormi-Nouri, R., Nilsson, L.-G., & Ohta, N. (2005). The novelty effect: support for the Novelty-Encoding Hypothesis. *Scandinavian Journal of Psychology*, *46*(2), 133–143.

http://doi.org/10.1111/j.1467-9450.2005.00443.x

Krendl, K. A., & Broihier, M. (1992). Student responses to computers: A longitudinal study. *Journal of Educational Computing Research*, *8*(2), 215–227.

Kulik, J. A. (1994). Meta-analytic studies of findings on computer-based instruction. In E. L. Baker & H. F. O'Neil (Eds.), *Technology assessment in education and training* (pp. 9–33). Hillsdale, NJ: Erlbaum.

Landsberger, H. A. (1958). *Hawthorne revisited: Management and the worker: its critics, and developments in human relations in industry*.

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and Evaluation of a User Experience Questionnaire. *HCI and Usability for Education and Work*, 63–76. http://doi.org/10.1007/978-3-540-89350-9_6

Lösch, E., Alt, F., & Koch, M. (2017). Mirror, Mirror on the Wall: Attracting Passers-by to Public Touch Displays With User Representations. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces* (pp. 22–31). ACM Press. http://doi.org/10.1145/3132272.3134129

Memarovic, N., Clinch, S., & Alt, F. (2015). Understanding Display Blindness in Future Display Deployments. In *Proceedings of the 4th International Symposium on Pervasive Displays (PerDis '15)*. ACM Press. http://doi.org/10.1145/2757710.2757719

Merikivi, J., Nguyen, D., & Tuunainen, V. K. (2016). Understanding Perceived Enjoyment in Mobile Game Context. In *Proc. 49th Hawaii International Conference on System Sciences*.

Müller, J., & Stocker, A. (2012). *Siemens Building Technologies Division: Globaler Wissens- und Erfahrungsaustausch mit References+*. Enterprise 2.0 Fallstudiennetzwerk.

Novelty Effect. (2017). Retrieved November 1, 2017, from https://en.wikipedia.org/wiki/Novelty_effect

Ojala, T., Kostakos, V., Kukka, H., Heikkinen, T., Lindén, T., Jurmu, M., … Zanni, D. (2012). Multipurpose Interactive Public Displays in the Wild: Three Years Later. *IEEE Computer*, *45*(5), 42–49. http://doi.org/10.1109/MC.2012.115

Ojala, T., Valkama, V., Kukka, H., Heikkinen, T., Lindén, T., Jurmu, M., … Hosio, S. (2010). UBI-hotspots : Sustainable Ecosystem Infrastructure for Real World Urban Computing Research and Business. In *Proceedings of the International Conference on Management of Emergent Digital EcoSystems (MEDES '10)* (pp. 196–202). http://doi.org/10.1145/1936254.1936288

Ott, F. (2018). *CommunityMirrors: Interaktive Großbildschirme als ubiquitäre Natural User Interfaces für Kooperationssysteme*. Universität der Bundeswehr München. http://doi.org/10.18726/2018_1

Ott, F., & Koch, M. (2012). Social Software Beyond the Desktop — Ambient Awareness and Ubiquitous Activity Streaming. *It - Information Technology*, *54*(5), 243–252. http://doi.org/10.1524/itit.2012.0687

Poppenk, J., Köhler, S., & Moscovitch, M. (2010). Revisiting the novelty effect: When familiarity, not novelty, enhances memory. *Journal of Experimental Psychology Learning Memory and Cognition*, *36*(5), 1321–1330. http://doi.org/10.1037/a0019900

Prochaska, J. O., & DiClemente, C. (1982). Trans-Theoretical Therapy - Toward a More Integrative Model of Change. *Psychotherapy Theory Research & Practice*, *19*(3), 276–288. http://doi.org/10.1037/h0088437

Schwarzer, J., Draheim, S., von Luck, K., Wang, Q., Casaseca, P., & Grecos, C. (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*. http://doi.org/10.1145/2971485.2971493

Schwarzer, J., & von Luck, K. (2012). Periphere Wahrnehmung von Unternehmensinformationen – ein Erfahrungsbericht. In J. Cleve (Ed.), *Proc. 8th Wismarer Wirtschaftsinformatik-Tage 2012* (pp. 82–91).

Siek, K. A., Hayes, G. R., Newman, M. W., & Tang, J. C. (2014). Field Deployments: Knowing

from Using in Context. In J. S. Olson & W. A. Kellogg (Eds.), *Ways of Knowing in HCI* (pp. 119–142). New York: Springer. http://doi.org/10.1007/978-1-4939-0378-8_6

ten Koppel, M. (2011). *The Impact of Display Form Factor on User Behaviour with Interactive Public Displays*. TU Berlin.

ten Koppel, M., Bailly, G., Müller, J., Walter, R., Behavior, P., & Koppel, M. (2012). Chained Displays: Configurations of Public Displays can be used to influence Actor-, Audience- and Passer-By Behaviour. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)* (pp. 317–326). Austin: ACM Press.

Thompson, C. (2014). The Novelty Effect. Retrieved November 1, 2017, from https://medium.com/message/the-novelty-effect-cf606715ae62

Tulving, E., & Kroll, N. (1995). Novelty assessment in the brain and long-term memory encoding. *Psychonomic Bulletin & Review*, *2*(3), 387–390. http://doi.org/10.3758/BF03210977

Webster, J., & Ho, H. (1997). Audience Engagement in Multimedia Presentations. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, *28*(2), 63–77. http://doi.org/10.1145/264701.264706

Nauwerck, G., Cowen Forssell, R. (2018): The Digital Work Environment–a Challenge and an Opportunity for CSCW. 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/ecscw2018_4

The Digital Work Environment–a Challenge and an Opportunity for CSCW

Gerolf Nauwerck
The Health, Technology and Organisations research group, Department of Information Technology, Uppsala University, Sweden.
gerolf.nauwerck@it.uu.se

Rebecka Cowen Forssell
Centre for Work Life Studies and Evaluation (CTA), Department of Urban Studies, Malmö University, Sweden.
rebecka.forssell@mau.se

Abstract. In this exploratory paper we will present the emerging concept of the Digital Work Environment. This is concept rooted in Swedish debate on the workplace, information and communication technology (ICT) and well-being. We argue that the concept can be understood as a boundary object uniting different actors (mainly researchers, unions, and policy makers) in a common discourse on what has been labelled as the dark side of information technology. We also argue that the concept needs to embrace an organisational perspective as well as the relational aspects of the psychosocial work environment. Such a move would open the door to a large volume of relevant research that might reinvigorate the concept. More specifically we will show how this would allow the inclusion of the increasingly important aspect of cyberbullying, which at the same time is an example of blurring borders between work and non-work ICT use.

Introduction

Walsham (2012) has been making the case for a more critical and socially oriented approach to research on information and communication technology (ICT), asking the question: "are we making a better world with ICT?" (Ibid. p. 91). Given the worldwide rise of workplace burnout (Carod-Artal & Vázquez-Cabrera, 2013) the workplace remains a relevant setting for researchers wishing to make a better world. Over the last decades enterprise ICT has become increasingly complex, as new generations of solutions are layered over the old ones. We still rely on ERP systems and legacy systems based on database technology (Armstrong et al. 2012, Bergin & Haigh, 2009), big data is all the rage (Walker, 2015) as is gamification (Robson et al., 2015), the Internet of Things (Lee & Lee, 2015), machine learning (Jordan & Mitchell, 2015) and so on. What

awaits around the corner we can only speculate in (cf. Neely, 2013). However, all this complexity not only brings benefits to organisations and employees, there are challenges too. Tarafdar et al. (2013) refer to this as the dark side of ICT. Writing on information and attention, van Knippenberg et al. (2015) summarizes the challenges on various levels:

> With these new opportunities for creating and capturing value, though, come pathologies for individuals, teams, and the organizations themselves. [—-] The pathologies that can result from such challenges run the gamut from exhaustion and burnout to impaired judgment, suboptimal decision making, wasted effort, and reduced productivity. (Ibid. p. 650)

The field of computer supported collaborative work (CSCW) has a long tradition of studying the relation between work and ICT. Regardless of whether CSCW in understood the strict sense (Schmidt, 2001) or in a wider sense (Blomberg and Karasti, 2013), it will have an impact on the Digital Work Environment. Hence, being aware of this impact should be relevant to the study of CSCW. Furthermore, as there is comparatively little research on the dark side of ICT (Tarafdar, 2013) CSCW researchers should also be able contribute towards filling a research gap in this highly important research area.

In this exploratory paper we reflect upon the concept of the *Digital Work Environment*, a concept strongly rooted in the traditions of the welfare state in Scandinavia in general and in Sweden in particular. As such, it might be hard to understand outside of this context. We will however argue that the concept has been successful in raising public awareness in Sweden on the dark side of ICT. At the same time we note that the current understanding of the concept is based on cognitivist thinking and a limited understanding of the social embeddedness of software. By introducing the example of cyberbullying we wish to point out that what constitutes the Digital Work Environment is constantly changing. So far, cyberbullying has received little attention in the discussion on Digital Work Environment. Still we will argue, that with the emergence of cyberbullying in working life new situations occur that forces us to widen the understanding of the Digital Work Environment.

The rest of the paper is structured as follows. First we take a closer look at the concept of the Digital Work Environment, its genesis, strengths and limitations. We then present cyberbullying as an example of a relevant problem that needs to be included in the discourse on the Digital Work Environment. We end with suggesting that the strength of the concept lies in its functioning as a boundary object and challenge the CSCW community to engage in the discourse.

The Digital Work Environment

We wish to point the CSCW community to the relative success of the emergent concept *Digital Work Environment* in Swedish public debate. The concept Digital Work Environment dates back at least to the year 2000, but it was with the release

of a polemic book by usability expert Jonas Söderström (2010) that the concept became part of mainstream debate. The book argues that much of workplace ICT –as a result of bad design and implementation–does not support the user, rather it often contributes to stress. Today the concept is used in Swedish media, in Swedish government reports, by Swedish unions and even in campaigns by Swedish telecom companies. A quick search for the term in the Retriever Research database (January 2018) yielded 351 hits from various Swedish news sources. The increasing popularity was also obvious, with 80% articles being published over the last four years (2014-2017).

The concept has it roots in the term *work environment*, which in turn is related to occupational health and safety. Thus, when the concept Digital Work Environment emerged in the Swedish public debate it was as a rhetorical figure alluding to established concepts such as the psychosocial work environment or the physiological work environment. (For an overview of the concept, see Abrahamsson & Johansson, 2013.) More specifically it is related to the term *cognitive* work environment (Lind et al., 1991). This concept is part of a longer research tradition including classic works such as *User centered system design* by Norman & Draper (1986) and *Cognitive Work Analysis* by Vincente (1999).

Related concepts such as technostress focuses on pathologies at the individual level, while causes remain a complex of antecedents. Thus, technostress denotes an individual problem. In contrast, the Digital Work Environment moves focus from the individual to the antecedents, in turn indicating that these can be addressed. Even the antecedents can however be understood as contextual or social, rather than just related to a single system. An example of this is the concept of technological gaps (Bailey et al., 2010). Technological gaps relates to technology interdependence. This can be understood as a specific aspect of coordination and (task) interdependence within organisations, where the technology itself brings certain constraints as well as strategies for closing the gaps.

There is no consensus on the definition of the concept. On the contrary, there is a continuum, ranging from what could be understood as a strong definition to an all encompassing, weak definition. The strong definition focuses on interface design and cognitive aspects, whereas the weak definition seeks to cover most aspects of the modern workplace:

> The work environment, with its problems and opportunities of physical, organizational, social and cognitive nature, which results from the digitization of work support systems and tools. (Gulliksen et al., in press)

It should also be noted that the concept up until now has not been used in a strict academic context and in the debate it is sometimes also avoided for the very same reason. Instead, the concept has been replaced with constructs such as “the impact of digitalisation on the work environment”. We will not comment further on the above definition here, instead we will focus on some limitations that are not the

result of the concept as such but rather the research tradition from which it originated.

Limitations to the Current Approach to the Digital Work Environment

Some researchers have claimed that research on the Digital Work Environment is limited (Gulliksen et al., 2015). This is however a rather strong claim. On the contrary, there can be no doubt that many of the issues related to the concept have been at the centre of research for quite a long time. Indeed, a large number of issues are identified as well as analysed and theorized in Zuboff's 1988 classic In the Age of the Smart Machine. Taking a few cues from Foucault, her idea of automating/informating remains a very powerful tool for analysis of the digital transformation and its effect on the Digital Work Environment.

We argue that this perceived research gap is the result of the research traditions in human computer interaction and ergonomics, where the concept is perhaps most frequently used. Taking a broader view will reveal much relevant research; this in turn may open up a whole new palette of concepts and frameworks that can be used to revitalize the research related to the Digital Work Environment. By this we do not mean that current research should be abandoned, rather that the field is too important to be dominated by just one tradition.

The following case can serve as an example of the difficulties that might arise from approaching software (in the workplace) in the same manner as hardware (in the laboratory).

One of the better-known stories relating to ICT and the work environment is the establishment of the TCO Certification of IT products (TCO Certified), based on life cycle criteria for social and environmental responsibility. This was the result of collaborations between researchers and unions and quickly went from a Swedish example to a global certification. As researchers continued to study the effects of ICT in the workplace it seemed like a good idea to replicate the success of the TCO Certification of hardware with a similar certification of software. Thus, the UsersAward project was initiated (Walldius et al., 2009). There were some significant differences in approach, one important one being that while the hardware certification was based on expert analysis, the UsersAward was to be based on users experiences of software in use (in line with both the ISO standard of usability and research). While the project resulted in some software certifications, the project did not experience anything similar to the success of the TCO Certification. It has been argued that this can be related to different methodologies for assessment:

> In the end it was no economic viability in combining TCO Certified with USER Certified. The former being an expert based certification of hardware related to environmentally and socially

> sustainable production, use and recycling and the latter a user based certification of software quality (Walldius et al., 2016, p. 138, our translation.)

Rather than attributing the difficulties to methodological issues, we put forward that it perhaps more related to the discrete nature of hardware versus the continuous nature of software. This is a similar argument to that raised by authors such as Ensmenger (2012) and Mahoney (2008) when discussing software history. Ensmenger stresses the differences in the following way:

> Whereas the computer itself was a definite material artifact that could readily be identified and isolated for testing, evaluation, and improvement, software systems were inextricably intertwined with a larger system of computing that included not just machines, but also people and processes. (Ibid. p. 762)

In short, hardware lends itself to certification quite easily, but software–especially enterprise software–does not. The reason for this being its embeddedness in a complex social setting, in Mahoney's words:

> Thus, the models and tools that constitute software reflect the histories of the communities that created them and cannot be understood without knowledge of those histories, which extend beyond computers and computing to encompass the full range of human activities. All software, even the most current, is in that sense "legacy" software. (Mahoney, 2008, p. 8)

Of course, we are not arguing that the researchers behind UsersAwards are not aware of this, on the contrary this is part of why they wished the certification to be based on users opinions. Still, there seems to be an underlying idea that both hardware and software are discrete artefacts. Another way of putting this would be to say that the material turn is being underway but not completed (cf. D'Adderio, 2010).

This, we believe, can in part be attributed to the genesis of the concept of the Digital Work Environment. This can be traced back to a concern for workers safety and health in the 1970s, which relied strongly on the field of ergonomics and human factors. This wide, interdisciplinary field has been described as having three focus areas (in order of importance): physical ergonomics, cognitive ergonomics and organisational ergonomics. (To make distinctions even less clear, Swedish legislation introduces the overlapping concepts of the physical and the organisational and social work environment.) Basic questions relating to the Digital Work Environment are still based on physical ergonomics, such as posture and mouse movements. Most (traditional) usability issues fall under cognitive ergonomics. The subfield of organisational ergonomics is in comparison fairly immature, yet it is–presumably–here that we would find an understanding of the more complex interplay between software and organisation including such phenomena as cyberbullying.

Software can be understood as being co-created in actual use, but in earlier decades there were still limited instances of software in a particular workplace and hence the effects were also more limited. This has radically changed now, with what can be summarized as a proliferation of ubiquitous computing. Thus,

the need to understand the social embeddedness is more critical today and this is why we believe that the concept of the Digital Work Environment needs to be updated with insights from research on contemporary working life.

Cyberbullying as an Emerging Issue in the Digital Work Environment

One relatively new issue in the concept of Digital Work Environment is the emerging phenomenon of cyberbullying in working life. Broadly defined, cyberbullying refers to deviant and hostile behaviour that involves the use of email, text messages, blogs and social network sites e.g. Facebook, or other information communication technologies (Kowalski, Limber, Agatston, 2012; Patchin & Hinduja, 2006). In line with most definitions of workplace bullying, cyberbullying creates situations where the target feels helpless and defenceless from the negative acts (Vranjes, Baillien, Vandebosch, Erreygers, & De Witte, 2017). Being a new phenomenon, cyberbullying in working life is so far an under-researched area (Bartlett, 2011; Göransson, 2011; Privitera, 2009; Brack, 2014). However, there is an increased volume of studies on cyberbullying in working life that show negative implications related to the targeted individual's health and job satisfaction (Coyne, 2016; D'cruz, 2013; Muhonen et al., 2017; Snyman, 2015). While research on cyberbullying in working life is still in its infancy, the youth literature provides complemented insight into how cyberbullying can be expressed. Kowalski et al. (2012) outlines a number of behaviours that they argue constitutes cyberbullying, such as; flaming (e.g. brief, heated exchange between two or more people, often on public online forum), cyber harassment (e.g. repetitive and persistent negative behaviour online directed to a specific target), denigration (e.g. spreading of false or cruel statements about another person online), ostracism (e.g. social exclusion of another person on password-protected online forums) and cyberstalking (e.g. the use of electronic communication to stalk another person).

Cyberbullying can be understood as the result of increased communication based on digital media. Email is the most commonly used communication technology in working life and two of three Swedish employees use email on a daily basis (Findahl, 2012). Enhanced by mobile devices such as smartphones and laptops, communication in working life has become increasingly less time and spatial bounded. Communication technology enables people to work and communicate from other places than the workplace and at other times than during work hours. As a consequence, previously separated boundaries between work and non-work have become extensively blurred. Compared to traditional workplace bullying, cyberbullying can continue or even begin when the working day is over. Thus, cyberbullying challenges previous understandings of *when* the bullying is work related. For managers who are obliged to act upon and prevent

bullying in the workplace, understanding the distinction between what is work and private related conflicts become crucial. Moreover, the accessibility in digital communication also means that new types of actors, not necessarily members of their own organisation, but clients, customers, students or pupils can with more ease than before target an employee. Thus, the communication technologies used in cyberbullying challenges traditional understandings of *who* are perpetrators in work life bullying. Understanding what type of relations are present in situations of bullying also has consequences for how these situations can be dealt with within a work organisation.

While most digital communication in working life is conducted via email, social network sites such as Facebook are becoming increasingly commonplace in work organisations. In the literature on organisational ICT, the use of social network sites is often divided into two main types of use. The use of public social networking sites for communication with external parties such as customers, vendors and the public at large, and the use of internal social networking sites owned by the organisation for internal communication only (Leonardi, Huysman, & Steinfield, 2013; Rooksby et al., 2009). Although social network sites have a widespread use for organisations, most social network sites are primarily used by individuals for private purposes. Nevertheless, when users are including connections from different spheres of their lives, including professional contacts, social network sites such as Facebook also become a platform for work relations (cf. Marwick, 2011; Vitak, Lampe, Gray, & Ellison, 2012). The use of private social network sites as platforms for the maintenance of work relationships further contributes to the perception of blurred boundaries between work and non-work, as well as the private and the professional.

Cyberbullying is an example of the complexity involved in the concept of Digital Work Environment. As cyberbullying cannot be limited to the physical workplace or to the relations associated within the workplace, the phenomenon of cyberbullying stresses how Digital Work Environment extends to involve places, situations and relations outside the physical workplace. Moreover, as cyberbullying in working life involve the use of private social network sites such as Facebook, we argue that cyberbullying challenges traditional understandings of what constitutes the workplace IT. On the one hand, social network sites are increasingly used within organisations as a tool for information dissemination, recruiting and promotion of the work organization (Vitak et al., 2012). Thus, social network sites can be understood as 'Shadow IT' of an organization i.e. as an unsanctioned supplement to the work organisation's IT portfolio (Rentrop & Zimmermann, 2012). On the other hand, a social network sites such as Facebook is primarily a platform for social interactions. As interaction on such sites not only includes co-workers but most often also family members and friends, employers' interferences may be a sensitive issue that generates concern regarding workers privacy and integrity.

The Digital Work Environment as a Boundary Object

There might seem to be a paradox here. On the one hand, we have argued that some of the shortcomings of previous efforts to address issues in the Digital Work Environment can be traced back to a conflation of hardware and software. On the other hand, we have also argued that in spite of these shortcomings the concept of the Digital Work Environment has proved successful in the context of public policy and debate. So, the question is how can the same concept be subject to a critique while at the same time seem so successful in public debate?

In reply to this we would like to draw on Star (2010) and her idea of boundary objects (for a discussion see Lee, 2007). In fact, we would like to suggest that the concept of the Digital Work Environment functions similar to a boundary object. Star argues that a boundary object allows "different groups to work together without consensus" (Star, 2010, p. 602), one important aspect of this being the boundary object's interpretative flexibility. The term has been used to discuss theoretical concepts such as resilience (Brand & Jax, 2007). This is in line with Star's own reasoning as she points out that a boundary object derives its materiality from being acted toward and with, not the other way around. Some researchers have also combined boundary objects in combination with communities of practice.

All in all, this provides a framework for understanding the concept of the Digital Work Environment. The loosely defined concept has interpretative flexibility and seems to allow researchers, unions and policy makers and even practitioners to engage in a common discourse over professional borders. We illustrate our argument with Figure 1, below.

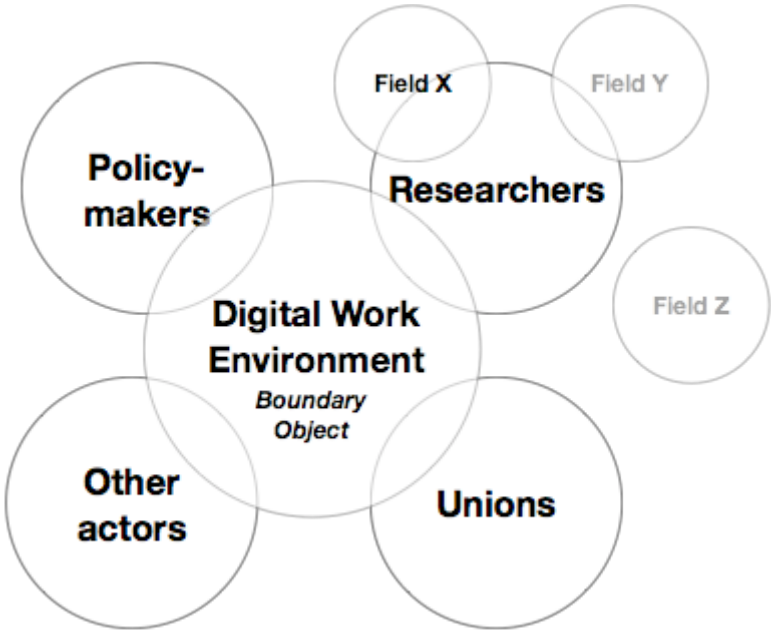

Figure 1. The Digital Work Environment as a boundary object.

The concept of the Digital Work Environment is here suggested to function as a boundary object. At the same time the research discourse has been dominated by the cognitive research tradition, and we argue that other research fields needs to engage in the same discourse.

It can also be argued that there is a need for flexibility and interpretation in the concept in a time where technological advancements emerge with a rapid speed (Rosa, 2014). Thus, what constitutes the Digital Work Environment is constantly negotiated and developed. Moreover, in order to explore the concept we must also pay notice to what constitutes the organisation's work environment. New behaviours and patterns emerging from communication technologies can be argued to challenge traditional understandings of the organisation's work environment and the workplace IT.

Discussion

To move forward, it is now necessary to build more strongly upon insights and research from the field of organisations and ICT. So far, the phenomenon of cyberbullying has received little attention in the discussion on the Digital Work Environment. Yet, the emergence of cyberbullying creates new situations that need to be handled on an individual as well as on an organisational level. Essential to this transformation of working life conditions and practices, we believe, are the introduction of digital platforms such as social network sites that tends to blur boundaries between the private and the professional. Thus, characteristic for cyberbullying is that the bullying behaviour cannot be restricted to the physical workplace or to the relationships within the workplace. Moreover, new platforms for work life interaction are constantly emerging. In other words, the example of cyberbullying in working life points to the limits of a cognitive approach to the analysis of the digital work environment. Thus, while cognitive aspects remain important it is also necessary to include organisational and relational aspect of the work environment. Finally it is also important to include technologies not traditionally associated with workplace ICT in the analysis, such as social networking.

The need for a wider approach is even more important as organisations themselves are changing rapidly, and we need to account for emerging forms, such as post-bureaucratic organisations (Kellog et al., 2006)–and phenomena such as the "gig economy" (de Stefano, 2016). An important part of this understanding comes, we would argue, from letting go of the idea of software as discrete artefacts and embracing the complexity that comes with an understanding of software as something inherently integrated in the practices of an organisation. An argument more eloquently stated by Orlikowski & Yates (2016):

> We believe that such approaches are particularly valuable as they afford the possibility of accounting for the messy, dynamic, contested, contingent, negotiated, improvised, heterogeneous, and multilevel character of ICTs in organizations. (Ibid. p. 132)

One important aspect of this is that if we wish to reduce the pathologies, the managers' attention is necessary. Indeed, creating organizational resilience relating to the Digital Work Environment can be understood as a capacity building process, working from the microfoundations and up (Eggers & Kaplan, 2013). Of course, taking this step also opens up to a complexity where there are no clear cut answers and different approaches emphasise different aspects, as shown in an overview of various strands of social constructivism, by Leonardi and Barley (2010). Nevertheless, as argued by Volkoff & Strong (2013) in relation to critical realism, there is value for practitioners as well as researchers in these approaches: to ensure that an organization achieves useful outcomes from IT" (Ibid. p. 832)–or in other words a good Digital Work Environment.

This call for a wider approach is in itself nothing new, in many ways it mirrors the argument in Bannon (1995) and the call to shift focus from human factors to human actors. Still, it seems evident that this call might need to be repeated. However, as the field of CSCW has embraced the practice approach (Kuutti & Bannon, 2014), we believe it is well positioned to develop our understanding of the Digital Work Environment. Furthermore, Wolf et al. (2015) claim that progress in the field of CSCW has "not been accompanied by wholesale acceptance in the commercial and industrial world" (Ibid. p. 2). Framing research as related to the Digital Work Environment should possibly make it more relevant to a larger community outside of academia, as made evident by the Swedish example.

Conclusions

In this paper we have discussed the emergent concept of the Digital Work Environment. We have argued that this concept has been successful in the Swedish policy debate, shining a light on the dark side of information technology. We claim that this is in part because the concept works as a boundary object. At the same time we stress that the concept is still strongly rooted in a cognitive tradition. Cognitive ergonomics are necessary but not sufficient approach to understand new challenges to the digital workplace such issues as cyberbullying.

Clearly, more analytical development is needed but we believe that the CSCW community is well positioned to contribute to research on the Digital Work Environment. Furthermore, finding a similar term would not only be a bridge between researchers but also between researchers and the public. Thus we end with a challenge to the community to find a suitable term–or to embrace the Digital Work Environment.

Acknowledgements

The authors would like to thank anonymous reviewers for very constructive suggestions. This being an exploratory paper, some of these will be addressed in later revisions.

References

Abrahamsson, L., & Johansson, J. (2013). One Hundred Years of Inertia: An Exposé of the Concept of the Psychosocial Work Environment in Swedish Policy and Research. *Nordic Journal of Working Life Studies*, *3*(1), 4–30.

Armstrong, S., Sandberg, A., & Bostrom, N. (2012). Thinking Inside the Box: Controlling and Using an Oracle AI. *Minds and Machines*, 22(4), 299–324.

Bailey, D. E., Leonardi, P. M., & Chong, J. (2010). Minding the Gaps: Understanding Technology Interdependence and Coordination in Knowledge Work. *Organization Science*, 21(3), 713–730.

Bannon, L. J. (1991). From Human Factors to Human Actors: The Role of Psychology and Human-Computer Interaction Studies in System Design. In J. Greenbaum & M. Kyng (Eds.), Design at work: Cooperative design of computer systems. L. Erlbaum Associates Inc.

Bartlett, J. E., & Bartlett, M. E. (2011). Workplace bullying: An integrative literature review. *Advances in Developing Human Resources, 13*(1), 69-84.

Bergin, T. J., & Haigh, T. (2009). The commercialization of database management systems, 1969–1983. *IEEE Annals of the History of Computing*, 31(4), 26–41.

Blomberg, J., & Karasti, H. (2013). Reflections on 25 Years of Ethnography in CSCW. *Computer Supported Cooperative Work (CSCW)*, *22*(4-6), 373–423.

Brack, K., & Caltabiano, N. (2014). Cyberbullying and self-esteem in Australian adults. *Cyberpsychology, 8*(2).

Brand, F. S., & Jax, K. (2007). Focusing the Meaning(s) of Resilience: Resilience as a Descriptive Concept and a Boundary Object. *Ecology and Society*, 12(1).

Carod-Artal, F. J., & Vázquez-Cabrera, C. (2013). Burnout Syndrome in an International Setting. In S. Bährer-Kohler (Ed.), Burnout for Experts: Prevention in the Context of Living and Working (pp. 15–35). Boston, MA: Springer US.

Coyne, I., Farley, S., Axtell, C., Sprigg, C., Best, L., & Kwok, O. (2017). Understanding the relationship between experiencing workplace cyberbullying, employee mental strain and job satisfaction: a dysempowerment approach. *The International Journal of Human Resource Management, 28*(7), 945-972.

D'Adderio, L. (2010). Artifacts at the centre of routines: performing the material turn in routines theory. *Journal of Institutional Economics*, 7(02), 197–230.

D'cruz, P., & Noronha, E. (2013). Navigating the extended reach: Target experiences of cyberbullying at work. *Information and Organization, 23*(4), 324-343.

De Stefano, V. (2016). The rise of the "just-in-time workforce": On-demand work, crowdwork, and labor protection in the "gig-economy". *Comparative Labor Law & Policy Journal*, 37(3), 471.

Eggers, J. P., & Kaplan, S. (2013). Cognition and Capabilities: A Multi-Level Perspective. *Academy of Management Annals*, 7(1), 295–340.

Ensmenger, N. (2012). The Digital Construction of Technology: Rethinking the History of Computers in Society. *Technology and Culture*, 53(4), 753–776.

Findahl, O. (2013). Svenskarna och internet 2012: Göteborg: SE (Stiftelsen för internetinfrastruktur).

Gulliksen, J., Lantz, A., Walldius, Å., Sandblad, B., & Åborg, C. (2015). Digital arbetsmiljö. Stockholm: Arbetsmiljöverket.

Gulliksen, J., Lantz, A., Sandblad, B., Walldius, Å., & Åborg, C. (in press). Digitaliseringen och arbetsmiljön. Studentlitteratur.

Göransson, S., Näswall, K., & Sverke, M. (2011). Psykologiska perspektiv på hot och våld i arbetslivet: Kunskapsöversikt.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260.

Kellogg, K. C., Orlikowski, W. J., & Yates, J. (2006). Life in the Trading Zone: Structuring Coordination Across Boundaries in Postbureaucratic Organizations. *Organization Science*, 17(1), 22–44.

Kowalski, R. M., Limber, S., Limber, S. P., & Agatston, P. W. (2012). *Cyberbullying: Bullying in the digital age*: John Wiley & Sons.

van Knippenberg, D., Dahlander, L., Haas, M. R., & George, G. (2015). Information, Attention, and Decision Making. *Academy of Management Journal*, 58(3), 649–657.

Kuutti, K., & Bannon, L. J. (2014). *The turn to practice in HCI* (pp. 3543–3552). Presented at the the 32nd annual ACM conference, New York, New York, USA: ACM Press.

Lee, C. P. (2007). Boundary Negotiating Artifacts: Unbinding the Routine of Boundary Objects and Embracing Chaos in Collaborative Work. *Computer Supported Cooperative Work (CSCW)*, 16(3), 307–339.

Lee, I., & Lee, K. (2015). The Internet of Things (IoT): Applications, investments, and challenges for enterprises. *Business Horizons*, 58(4), 431–440.

Leonardi, P. M., & Barley, S. R. (2010). What's Under Construction Here? Social Action, Materiality, and Power in Constructivist Studies of Technology and Organizing. *Academy of Management Annals*, 4(1), 1–51.

Leonardi, P. M., Huysman, M., & Steinfield, C. (2013). Enterprise social media: Definition, history, and prospects for the study of social technologies in organizations. *Journal of Computer-Mediated Communication, 19*(1), 1-19.

Lind, M., Nygren, E., Sandblad, B., 1991. Kognitiva arbetsmiljöproblem och gränssnittsdesign. (In Swedish: Cognitive Work Environment Problems and Design of User Interfaces) CMD Report 20/91, Center for Human–Computer Studies, Uppsala University, Sweden.

Mahoney, M. S. (2008). What Makes the History of Software Hard. *IEEE Annals of the History of Computing*, 30(3), 8–18.

Marwick, A. E. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New media & society, 13*(1), 114-133.

Muhonen, T., Jönsson, S., & Bäckström, M. (2017). Consequences of cyberbullying behaviour in working life: The mediating roles of social support and social organisational climate. *International Journal of Workplace Health Management*, *10*(5), 376-390.

Neely, E. L. (2013). Machines and the Moral Community. *Philosophy & Technology*, 27(1), 97–111.

Norman, D. A., & Draper, S. W. (1986). User centered system design: New perspectives on human-computer interaction. Hillsdale, N.J: Erlbaum.

Orlikowski, W. J., & Yates, J. (2016). ICT and Organizational Change. *The Journal of Applied Behavioral Science*, 42(1), 127–134.

Patchin, J. W., & Hinduja, S. (2006). Bullies move beyond the schoolyard a preliminary look at cyberbullying. *Youth violence and juvenile justice, 4*(2), 148-169.

Privitera, C., & Campbell, M. A. (2009). Cyberbullying: The new face of workplace bullying? *CyberPsychology & Behavior, 12*(4), 395-400.

Rentrop, C., & Zimmermann, S. (2012). Shadow IT. *Management and Control of Unofficial IT. ICDS*, 98-102.

Robson, K., Plangger, K., Kietzmann, J. H., McCarthy, I., & Pitt, L. (2015). Is it all a game? Understanding the principles of gamification. *Business Horizons*, 58(4), 411–420.

Rooksby, J., Baxter, G., Cliff, D., Greenwood, D., Harvey, N., Kahn, A., . . . Sommerville, I. (2009). Social networking and the workplace. *The UK Large Scale Complex IT Systems Initiative*, 1-39.

Rosa, H. (2014). Acceleration, modernitet och identitet. Tre essäer: Göteborg: daidalos.

Schmidt, K. (2012). Editorial. *Computer Supported Cooperative Work (CSCW)*, 21(6), 475–484.

Snyman, R., & Loh, J. M. (2015). Cyberbullying at work: The mediating role of optimism between cyberbullying and job outcomes. *Computers in Human Behavior, 53*, 161-168.

Söderström, J. (2010). Jävla skitsystem: hur en usel digital arbetsmiljö stressar oss på jobbet - och hur vi kan ta tillbaka kontrollen. [Stupid **** system! Stress caused by a bad Digital Work Environment and how we can reclaim control.] Stockholm: Publit Sweden.

Star, S. L. (2010). This is Not a Boundary Object: Reflections on the Origin of a Concept. *Science, Technology, & Human Values*, *35*(5), 601–617.

Tarafdar, M., Gupta, A., & Turel, O. (2013). The dark side of information technology use. *Information Systems Journal*, 23(3), 269–275.

Vicente, K. J. (1999). Cognitive work analysis: Toward safe, productive, and healthy computer-based work. CRC Press.

Vitak, J., Lampe, C., Gray, R., & Ellison, N. B. (2012). *Why won't you be my Facebook friend?: Strategies for managing context collapse in the workplace*. Paper presented at the Proceedings of the 2012 iConference.

Vranjes, I., Baillien, E., Vandebosch, H., Erreygers, S., & De Witte, H. (2017). The dark side of working online: Towards a definition and an Emotion Reaction model of workplace cyberbullying. *Computers in Human Behavior, 69*, 324-334.

Volkoff, O., & Strong, D. M. (2013). Critical Realism and Affordances: Theorizing IT-associated Organizational Change Processes. *MIS Quarterly*, 37(3), 819–834.

Walldius, Å., Sundblad, Y., Bengtsson, L., Sandblad, B., & Gulliksen, J. (2009). User certification of workplace software: assessing both artefact and usage. *Behaviour & Information Technology*, 28(2), 101–120.

Walldius, A., Andersson, A., & Löfstrand, C. (2016). När digitaliseringen förvandlar arbetet och ekonomin – om fackets nygamla roll som granskare och pådrivare. In Sandberg, Å., På jakt efter framtidens arbete.

Walker, R. (2015). Impact of Analytics and Big Data on Corporate Culture and Recruitment. From Big Data to Big Profits (pp. 184–201). Oxford University Press.

Walsham, G. (2012). Are we making a better world with ICTs? Reflections on a future agenda for the IS field. *Journal of Information Technology*, *27*(2), 87–93.

Wulf, V., Schmidt, K., & Randall, D. (2015). Introduction: Meeting the Challenge of Change. In V. Wulf, Schmidt, K., & D. Randall (Eds.), *Designing Socially Embedded Technologies in the Real-World* (pp. 1–8). London: Springer London.

Zuboff, S. (1988). In the age of the smart machine: the future of work and power. New York: Basic Books.

Andersson, C., Cozza, M., Crevani, L. and Schunesson, J. (2018): Infrastructuring for remote night monitoring: frictions in the strive for transparency when digitalising care services. In: Proceedings of the 16th European Conference on Computer-Supported Cooperative Work - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2018_6

Infrastructuring for remote night monitoring: frictions in striving for transparency when digitalising care service

Andersson, Christoffer; Cozza, Michela; Crevani, Lucia; Schunnesson, Jonathan
Mälardalen University, Sweden
lucia.crevani@mdh.se

Abstract. The question of how to organise for the introduction of a new service involving the interaction of humans and technologies is both crucial and challenging. Convergence between the community of practice using the technology and the design of the technology is crucial for the technology to become meaningful and usable. While processes of convergence are challenging in themselves, they become more complex if several communities of practice are going to use and collaborate around/through the technology. The co-presence of different communities of practice is a common situation when delivering public welfare services. In particular, the development of welfare technology is a context rich in potential frictions, making convergence challenging. By mobilising the concept of transparency, we analyse the process of implementation of remote night monitoring and highlight how transparency is related to different aspects. Such analysis reveals that processes of convergence are related in this context not only to frictions shared with other settings, but also to specific frictions related to matters of concern in welfare services. This leads us to discuss whether digitalised care services can be argued as still having a human side or not.

Introduction

The question of how to organise for the introduction of a new service involving the interaction of humans and technologies has proved to be both crucial and challenging. As scholars from different disciplines have shown, technologies run the risk of never becoming part of the practices they were meant to improve (Bijker et al., 1987; Gherardi, 2010; Suchman, 1987). When they do, convergence between the community of practice (Lave and Wenger, 1991) using the technology and the technology itself can be observed. More specifically, we are talking of convergence in the sense of mutual alignment between practices – socially and materially sustained patterns of action with a normative character – and the emergent design of technology (Star et al., 2003). From this point of view, design is not finished before implementation (Aanestad, 2003).

While such processes of convergence are challenging in themselves, they become even more complex if several communities of practices are going to use and collaborate around/through the technology being introduced. As Star et al. (2003) show, for instance, this is a process characterised by politics and power enactment. The co-presence of different communities of practice is a common situation when delivering public welfare services (Breskovic et al, 2013; Cozza et al, 2016; MacManus et al., 2013). Such services are currently being changed by introducing new digital technologies (Östlund et al, 2015). This is particularly true in the case of what in Scandinavian are called welfare technologies, that is, technology used "to improve the services provided by the welfare society and make them more efficient" (The Nordic Centre for Welfare and Social Issues, 2010,p. 7). The issue of quality and efficiency of welfare services has been identified as one of the most pressing challenges to be addressed globally, given a rapidly ageing population. It is in fact estimated that by 2030, people over the age of 65 will represent about 24% of the population in Europe, 22% in the United States, and 12% in Asia and Latin America (Czaja & Schulz, 2006, p. 6). As anticipated, introducing technology in this context is challenging given that such technologies affect and are affected by the work of different communities of practice with a long history and strong professional norms, such as nurses, social workers, physiotherapists, care assistants, legal experts, and politicians, etc (Kylberg et al, 2015).

Furthermore, high expectations of both increased efficiency and quality of service, combined with political pressure to digitalise (both nationally and at local level) met by scepticism in the public debate around certain technologies, make such a context rich in potential frictions to be dealt with when introducing digital technologies. Such frictions may be related both to convergence between technology design and user's practices and to convergence between technology

design and work practices within the public organisation. In this paper, we limit our attention to what happens when the technology and the organisational practices need to become aligned in order for those involved in delivering the service to be able to do their work properly in collaboration. This focus is empirically driven as we observed how much work was needed in order to try to achieve such an alignment, something that has been explored only to some extent in the literature discussing computer-supported cooperative work in social services (with more focus on healthcare than care, see for instance Hartswood et al, 2003). The convergence between technology and user's practices will not be analysed in this paper – future work will focus on that.

In this paper we are therefore going to explore which frictions emerge when trying to make a new (in this context) technology and work practices converge in order to deliver a social care service. Such an analysis provides a contribution to the literature on welfare technology (Peine et al, 2015; Östlund et al, 2015) as it highlights important issues to take into consideration, but also to the literature on infrastructure and convergence (Star et al, 2003) as we will discuss whether the context of social services presents peculiar challenges.

To this end, we mobilise the concept of transparency as introduced by Star et al. (2003) for conceptualising the process through which a technology becomes invisible as it develops to become part of a (in turn developing) practice. Transparency would, in other words, be the ideal state if convergence was complete – while this may be not achievable, practitioners strive towards it. We focus on a particular case in one Swedish municipality in which a technology for monitoring (a camera substituting visits from the homecare personnel in person during nights) was undergoing implementation in order to qualify what transparency may be about in this context. The Swedish case is particularly interesting given that municipalities have to provide welfare services and the pressure to digitalise is very high (Søndergård, 2014). We thus show that transparency is related to different aspects. Such analysis reveals that processes of convergence are related in this context not only to frictions shared with other settings, but also to specific frictions related to matters of concern in welfare services. This leads us to discuss whether digitalised care services can be argued as still having a human side or not.

The paper is organised as follows: after a short section introducing the theoretical framework, we discuss our method and present the case; we then analyse the endeavour to achieve transparency in relation to different issues in the process that we followed and conclude with a short discussion on the implication of our analysis so far. Given that this is an exploratory paper, such conclusions

are preliminary and would greatly benefit from discussion in order to be developed.

Infrastructuring: The challenge of transparency

When new technologies are introduced, they become connected to other technologies and systems in order to function. In the field of information systems, the term infrastructure has been commonly used to denote the network of interconnected visible and invisible (to the user) devices operating according to standards (for a critique, see Pipek and Wulf, 2009). Such a view has been criticized for considering infrastructure "a thing" and a "neutral" thing (Star and Ruhleder, 1996). An alternative is to build on a more relational and processual ontology and to think of infrastructure as relations embedding choices and politics (ibid). In other words, no device is isolated and a discrete entity – the development and use of the single device emerges in complex relationships (ibid). Not only does the device become embedded in complex organisational processes, but there needs to be a convergence between how the device connected to the system works and the work practices that the organisation has developed, in order for the device to be used in a meaningful way (Pipek and Wulf, 2009; Star and Ruhleder, 1996). In other words, it is hard to a-priori structure and automate a task; also, if users do not use a technology this is not a question to be addressed as overcoming user's resistance, but an organisational and learning challenge (Star and Ruhleder, 1996).

One central feature of an infrastructure is thus its invisibility in use: people that come in contact with the infrastructure are not constantly aware of the infrastructure, if it has become integral to work practices (Pipek and Wulf, 2009). Hence, when introducing a new technology, one concern is whether it, and the infrastructure it is part of, will become invisible to the users. Such a challenge also has consequences for how the technology can be designed and introduced – we can expect an iterative process in which the technology and the work practices need to adjust in relation to each other.

The concept of transparency can be mobilised in order to explore such an iterative process. A technology or system is transparent if users do not need to bother with how the underlying machinery or software functions (Star et al., 2003). Transparency indicates that users can actually make use of the technology without having to be worried about how the technology works. With the relational approach that we subscribe to, transparency is not an objective feature of the technology, nor is it something that can be designed a-priori. Rather, transparency emerges as information resources and social practices become

aligned. Transparency is no stable feature either; it is always provisional and may require further effort as politics or knowledge, for instance, develop (Star et al., 2003; Pipek and Wulf, 2009; Star and Bowker, 2002).

Star et al. (2003) showed how transparency has to do with the convergence of information artefacts and communities of practice in the case of wider-scale information systems. We build on their work and explore how transparency is striven for when an existing technology is introduced in a new context. We say "striven for" as absolute transparency may be impossible to achieve, given the relational and processual nature of transparency and infrastructuring. As Bowker and Star point out

> Transparency is in theory the endpoint of the trajectory of naturalization, as complete legitimacy or centrality is the endpoint of the trajectory of membership in a community of practice. Due to the multiplicity of membership of all people, however, and the persistence of newcomers and strangers as well as the multiplicity of naturalization of objects, this is inherently non-existent in the real world […]". (2000, p 311)

As they suggest, we can explore the emergence of transparency by paying attention to the emerging convergence between practices of using the technology, and the technology and its infrastructure. Star et al. (2003) propose referring to the concept of community of practice in order to conceptualise the social component of such a process. Community of practice (Lave and Wenger, 1991) denotes a group of people that share and sustain certain sociomaterial practices (Orlikowski and Scott, 2008), that is, routinised ways of doing things, language, technologies and norms. The new (in the context) technology and its infrastructure converge with a community of practice as its use becomes aligned with its design and access – as they co-constitute each other (Star et al, 2003). Such a process of convergence is highly situated, political, possibly contested and enacting power. Moreover, if, as often, more communities of practices are involved, each with its own power enactments, the process of convergence becomes even more complex. Negotiations are required, and they involve the practices in use, but also other technological artefacts that are part of the infrastructure and the norms, priorities, values that they materialise. Convergence is thus "a layering of solutions and conventions, memberships and standards" (Star et al, 2003, p 263).

After having presented some methodological considerations and briefly described the case we focus on, we will explore which frictions concerning transparency within the public organisation seem to emerge as remote night monitoring is implemented and then discuss what this may imply when it comes to technology used in social services.

Method

The work presented in this paper is part of a project on welfare technologies and older people that we have been working with for two years (including a pre-study). In this project, we cooperate with two municipalities in Sweden with the aim of developing knowledge about how new technologies are introduced and fostering a discussion on how such processes may be improved. We have therefore worked with both older people (older than 65 years) and representatives of the municipalities in different ways in order to investigate different aspects of the introduction and use of technology when providing social services to older people. After an initial phase devoted to understanding the relationship between technology and older people's needs, we turned our attention to two cases of introduction of new technology. We discussed which case to follow with each municipality as we wanted to focus on something that was ongoing, as well as both relevant and challenging for them. In this dialogue, we decided to focus on the introduction of remote night monitoring at one of the municipalities. Such a technology has already been introduced in a few other cities, is seen as very promising by public actors, but is still controversial in the public debate. Such a combination of aspects enabled us to follow a process that we anticipated would be complex and potentially revealing frictions between communities of practice.

Given the exploratory nature of our work, in this paper we adopt a qualitative approach aimed at closely following what the practitioners called "the implementation project" as it unfolded (Czarniawska, 2008). Inspired by ethnographic work as mobilised in the study of organisations (Cooren et al, 2008; Pinkand Morgan, 2013), we produced empirical material by interviewing the project manager, by observing project meetings and by collecting printed material, both documents publicly available and internal documents. We were allowed to record both when interviewing and when observing the meetings, and therefore to closely read these conversations (Dubois and Gadde, 2002). In the first reading, we looked for the terms in which the new service was being constructed, the terms in which the technology was being discussed and the practices that needed to be connected in order for the new service to function. From this reading, the relevance of the concept of transparency became clear and we also observed that we could distinguish between different kinds of transparency. In a second phase, we thus organised our analysis around different kinds of transparency. The result is presented in the section "striving for transparency". First we introduce the studied process.

Introducing remote night monitoring for older people at AnyTown

The implementation process studied can be understood in light of the increasing concern in Swedish municipalities regarding the ageing population that leaves them with the challenge of providing welfare services to an increasing number of older people at the same time as they face the problem of fewer people being available for, and wanting to work with, care services.

Such developments are globally relevant, but it may be argued that the challenge in Sweden, and other Nordic countries, is particularly acute as municipalities are supposed to offer welfare services to citizens who have the right to be supported to a greater extent than in other countries (Hanson et al, 2011). This has led to several projects started in several municipalities in which different kinds of technological solutions are tested (Søndergård, 2014). One of the technologies that has received a lot of attention more recently is cameras installed in older people's homes to remotely monitor their well-being during night-time – while normally there would be care workers from the home-care unit coming in person and checking that the older person is sleeping and well a couple of times per night. Remote monitoring happens in a similar way, with people in a control room turning the camera on remotely a number of times per night (depending on the older person's needs) and checking that the person is still sleeping. Typically, this service is provided to older people living alone and afraid of falling in the night or of encountering other problems during the night.

Before such a service (in person or remotely) can be delivered, a social worker makes an assessment of the patient's eligibility. At AnyTown, when the studied implementation project ends, the idea is that the social worker, in dialogue with the older people, will decide whether the monitoring at night should be performed in person or remotely. After that, the information will be delivered to the unit that provides home care at night. They will contact a private company that provides the camera and the technology needed for streaming images from the older person's home. Technicians from the company will then install a small camera in one of the top corners in the patient's bedroom (as well as other appropriate places in their home if deemed necessary). These cameras are then remotely monitored centrally by home-care personnel who turn them on at the decided intervals. The intention is to use the camera as a technical aid to perform regular check-ups on older people, and the municipality is clear in its communication that this is not to be confused with surveillance cameras to observe or control the user. Ideally, this will not only free up time for the municipality personnel, but also be less of an annoyance to the users who can avoid being woken by regular in-person check-ups.

At AnyTown, we followed the project that started once the procurement phase was concluded (a company had won the contract for providing the technology) in order to implement this technology. AnyTown is a municipality located in Sweden with roughly 100 000 inhabitants; the municipality as an organisation has 10 000 employees. The organisation includes different areas and units; in social services, there is a unit that assesses people's needs (social workers), a unit that provides home-care during the day, a unit that provides home-care during the night (the Night Patrol), a unit taking care of safety alarms (Alarm group), a unit taking care of quality and IT systems, among others. The project group that worked with the studied implementation process included representatives from all these units, some of them being managers and others being co-workers. In addition, a communications officer was included. Naturally, the involvement and collaboration between the project members varied with time, but the group had monthly meetings throughout the entirety of the project to make sure things were going in the right direction, to inform each other on what was going on in the different units regarding the project, and to bring to the surface challenges and problems and address them. The project group also shared a workspace online with all the documents that they deemed relevant for the project and were otherwise in contact via email. In the following section, we present our preliminary analysis of what happened in this project with regard to transparency.

Striving for transparency – making different communities of practice and a technology converge

In the municipality we studied, the project aimed at implementing the camera encountered a number of challenges and unexpected turns. In this paper, we interpret such unfolding of events in terms of necessary steps in order for the new technology to become close to transparent to the municipality employees. While, given the results of previous studies (cf. Star et al, 2003), it is not surprising that there are challenges and conflicts to be dealt with before the technology can become transparent, in the case of home-care services, we have particular communities of practices involved, which makes the process of convergence different from other organizational settings. It should also be kept in mind that this technology has received a lot of attention both in the public sector and more generally in Sweden, given the media and their reports. Hence, the technology is expected to create worries and problems, both within the organisation and in the municipality, which makes its becoming transparent both difficult and urgent for the people working with its implementation.

It is also interesting that the technology used, a camera installed in a room and activated at regular points in time by personnel sitting in a control room, is already used by the company that provides it in other contexts, such as for the surveillance of warehouses.

In the following sections, we empirically explore transparency in relation to certain aspects; see a summary in table I. We select those aspects that emerged as important in the observed meetings. Of course, other aspects may have been discussed in other arenas, and our aim is not to exhaustively reproduce all possible aspects. Rather, by exploring these selected ones we think we can learn something about introducing technology in the context of social services.

Transparency in relation to eligibility for the new service

The new service is part of the social services provided by the municipality, which means that it is the citizens' right to apply for it. As it is also a service that has been described in critical terms in the public debate at times and a service that is supposed to be more resource-efficient than when night monitoring is performed by visiting the older people in person, discussions during the project emerge about how and when to offer the possibility to receive remote night monitoring. For instance, when the accounting unit performs an estimation of the cost of the new service and compares it with the cost of visits in person, they realise that there are big savings to be made. As part of enacting accounting practice, they also calculate the optimal "batch size" to be achieved in order to optimise resources. That is, they propose to gather x number of people interested in the camera and then to offer it to all of them and at the same time, thus making another resource free. This is received with disconcertion mixed with irony by the project group, and different professionals forcefully restate that the citizen is provided with the service as soon as a need for it is assessed.

Such a strong statement about what constitutes the mission of the public officer meeting the older person in need is also mobilised when discussing how to communicate about the new service and which people to target. The management, somewhat worried by how the service is perceived and received, seem to be keen to start with the "proper older people" in order to gain a smooth start and accumulate some success stories – their practices being embedded in work devoted to maintaining legitimacy for the unit/organisation. The professionals, on the other hand, mobilise ethical considerations regarding who is to decide which older persons should be targeted and which not. Both these instances of friction are resolved by enacting the ethos of the public service that should be impartial on the issue of who is eligible and only decide based on the actual needs.

Transparency in relation to the appropriateness of service

The service is meant to make the older person feel safe, which means that the question of what happens when the camera is not working or while waiting for the camera to be installed is a central one if the service is to be trusted by both the older person and the employees involved in the provision of the service. As the project group starts working, such questions are brought up for discussion. For instance, what happens if the camera is not working because of technical failure? How long should the personnel wait before taking the initiative to activate someone to provide the service in person? What is an appropriate service and who is to decide?

Also, according to the way in which the municipality works, once the social worker makes the decision to grant a service, the service needs to be activated directly as the practices are to be appropriate to the older person's need which is imminent. The company, on the other hand, needs five days from order to delivery for installation of the camera given how their practices are organized. This means that there are five days of "limbo" between the decision and the service being provided according to the decision. Such an issue emerged during the project and had not been addressed before.

Both frictions are solved by including the older person in defining the routines for providing the service in her/his particular case. The project group seems very satisfied with such a solution and the new technology will probably become quickly transparent in this respect since the staff do not need to be worried about how to handle it as the question is settled in the initial care planning meeting with the older person.

Transparency in relation to assuring personal integrity

Part of the work done in the implementation project is directed towards adjusting current practices in the company providing the technology and in different units at the municipality in order to align them with the rules and expectations of a social service that respects personal integrity.

One friction arises from the routines for handling information in the company being situated in a private sector context in which nonhumans are surveilled, while the service being implemented concerns humans. This places demands on the security of the systems and on how personal details are managed, given Swedish legislation and given routines at the municipality. At the same time, it is important that the right information is shared between the organisations at the right time. Such a friction is handled by mobilising technological solutions that need to be integrated with the existing systems for instance, a login system

provided by an independent app – but the integration of the systems turns out to be less smooth than anticipated. In addition, securing integrity does not allow the e-commerce system to be used for every transaction, which the procurement unit would like, and a paper system will also need to be in place in parallel with the digital system. Such a friction is also handled by resorting to legal documents specifying the emerging practice, even though the company puts up resistance to this.

As the project unfolds, this friction persists. While the municipality repeatedly expresses the wish to solve the issue – with our lenses making the technology transparent – as soon as possible and the project members seem to work on as if the issue had been solved – just to be surprised again and again that it has not –, the company does not succeed in aligning their routines with what is required by the service they want to provide.

Transparency in relation to IT operations

The new technology and the related IT system are introduced by means of a project and as the project develops, the question of who is to take care of the IT system emerges. As there is some time pressure, the task is assigned by the project team to one of the project team members who seems to possess the competence and show initiative. It is the person who is responsible for the staff monitoring the older people through the camera. On the other hand, it is also recognised that this is not a good solution in the long term. Such a discussion also intertwines with a more general discussion at the municipality regarding how to assign ownership and administrative duties for digital systems in general and which principle to use. As this system may be used not only for older people but also for people requiring assistance due to disability, the new policy points to the head of department as most appropriate person to assume this responsability and to the unit taking care of shared systems as those who will maintain the new system. Such a solution would need to be approved by the manager but there is no opposition in the project group to shifting these responsibilities (in other words, there is no friction between the particular system and the rest of the digital systems and the practices developed around their functioning). Convergence is thus aided by leaning on organisational policies assigning responsibilities, and transparency in relation to administration seems to be unproblematic at this point of time. Future observations will reveal whether unpredicted frictions may emerge when the responsibility is actually transferred.

Table I. Summary of important aspects related to frictions in striving for transparency.

Transparency in relation to	Friction in the convergence of	Caused by	Handled by resorting to	Development
Eligibility	Social service practices, accounting practices, management practices in relation to legitimacy	Different logics when it comes to how to offer the service (citizen right, efficiency, probability of success)	Social services ethos	The friction is resolved
Appropriateness of service	Social service practices, legal practices at the municipality, company work practices, an artefact (the camera) that needs to be introduced (while "normal" home care staff is already available)	Timeliness being treated differently by the organisations	Including the older person in establishing routines and defining what is appropriate	The friction is resolved
Personal integrity	Social services work practices, private surveillance work practices, e-commerce practices, different IT systems	Different context of use of the technology, different priorities (for instance efficiency vs security)	- More technology - Legal documents - Paper documents in parallel to the digital ones	The friction is treated as solved but persisting
IT operations	Practices related to the use of this specific technology, IT operations and maintenance in general	Different knowledge of the new service and technology	Organizational policies	The friction seems easy to resolve in the future

When convergence of practices and technology remains incomplete

In the previous section, we have described some frictions that seem resolved or possible to resolve. Sometimes we observe that the convergence needed in order for the new technology to become transparent in relation to one dimension is a challenge to convergence in relation to another dimension. It is impossible to say

whether such tensions are caused by the process we observed still being at the intial phase or if they denote conflicting demands that will not be solved.

One example is the issue of personal details and how such information needs to be handled. As there is no secure system for transferring personal details on the older person to which the service is provided from the municipality to the company and no secure way to guarantee that the transmission from the camera is not hacked, the project team decided to use pseudonyms in the form of a combination of letters and numbers in the system. Hence, rather than Maria Camacho, the company will know that CV3 needs a camera, they will register that a camera has been installed at CV3 and the people monitoring Maria Camacho with the camera will also read in the system that CV3 needs to be checked on three times a night, for instance. If a hacker captures the transmission from the camera, s/he will not know which person s/he is looking at. This is a solution adopted in order to ensure personal integrity and make technology transparent in relation to it. But, in the course of the project, such a solution gave rise to another worry, which was in particular brought up by the legal advisor enacting practices embedded in norms related to the liability of the organisation. What was discussed is that the use of pseudonyms could increase the risk of providing the wrong service to the older person (and what legal consequences this might have), as it is easier to type the wrong sequence of letters and numbers in the system when registering a user and/or not to realise if an error has been made given that the staff is dealing with pseudonyms rather than with names of people that they can link to a face and personal history. Hence, transparency with regard to personal integrity benefits from this arrangement of practices and technologies, but transparency in relation to the appropriateness of service is suffering from this solution and remains incomplete with people at the municipality wondering what might happen if an error is made and not discovered.

Discussion

Our analysis reinforces the argument that technology in use is situated: the camera already exists and is already used, for instance, in warehouses, but once it becomes part of the delivery of social services, its use and the issues related to its use change (Aanestad, 2003). In elevators or warehouses, cameras are already transparent, or close to transparent, to us as well as to those interacting with them in organisations. A camera in an older person's house is not the same artefact, is not part of the same kind of infrastructure and does not need to align with the same work practices (Bratteteig and Ina Wagner, 2013). Hence, understanding the technology in relation to work practice when designing it is important, but it is also crucial to consider "implementation" as a process in which both technology

and practice are adjusted and co-evolve, rather than expecting implementation to just be about a "finished device" to be put into use (Pipek and Wulf, 2009, Aanestad, 2003).

As it is possible to see in table I, which summarizes our findings, different resources are mobilised in order to increase convergence. Depending on the friction, we see discussions resorting to more technology, to formal documents, to the professional ethos, but also to involving users in order to resolve tensions in the emergent alignment of practice and technology. Hence, as infrastructuring is sociomaterial (Orlikowski and Scott, 2008), working with increased transparency also includes heterogeneous elements that are put in relation. The project manager in the case studied was receptive to such different elements and seemed to be able to navigate the process without getting bogged down in the technology only, or the work practices only. We observed no serious conflict in the studied case, but this is, of course, not the same as claiming that managing a project in which convergence seems to increase is the same as handling a democratic process. The accounting practices were, for instance, not allowed to influence how decisions were made, which is an example of enactment of power.

Based on our theoretical framework, we could also argue that one could consider implementation as being accomplished once the technology is transparent. As we have seen, given the processual and relational nature of infrastructures (Star and Ruhleder, 1996), complete transparency may be impossible to achieve and is only temporary. In the case of social services, not only does technology develop and new devices are added, but also norms, laws, professional standards, etc change, which means that there are many factors that will, most probably, make striving for transparency an ongoing process, something that also has consequences also for how new technologies are maintained (Pipek and Wulf, 2009).

If we look at the kind of aspects we have identified when analysing the endeavour to achieve transparency, we can see that some are common to other settings, such as convergence in relation to administration, while others could be more specific to social services, such as personal integrity or the appropriateness of service. Interestingly, we see a relation between these aspects and certain matters of concern that Crevani and Cozza have identified with respect to the relation between older people's needs and technology (2018) – matters of concern in the Latourian sense of matters that are "real" although not easily characterised (as matters of fact could be), matters that gather interest and affect, given their complexity and relevance (Latour, 2004). For instance, frictions in relation to personal integrity can be related to the matter of concern "balancing appropriate care and integrity", that is, the tension between providing care at the right level

(which means letting others come close to the individual in order to make decisions about when an intervention is needed) and respecting personal integrity. Also, frictions related to eligibility for the new service could be argued to be related to the matter of concern "balancing providing sustainable services against including all older people (services accessible to everyone)", that is, the tension between providing services that are possible to deliver and having to cope with the limited resources available.

Finally, what our analysis also shows is that all the work necessary to try to make the camera transparent indicates that providing home care through technology is not the same as forgetting what home care is about. The question is whether one could argue that there is still a "human side" to remote monitoring. A positive answer would be motivated by the frictions we see that the organisation needed to handle. Many of these frictions are related to the nature of the welfare service provided: care. Providing care requires certain practices that are still relevant even when technology is the actor present in the older person's home. A negative answer would be motivated by arguing that there is a difference if monitoring is performed by a human in person. Even though the user may sleep throughout the visits (whether they are done through the camera or in person), there is still a symbolic and affective value in knowing that a human has come and looked after the older person. This is lost if such visits in real life are substituted by a camera (Bratteteig and Ina Wagner, 2013). Although there is no simple answer to such questions, we believe that they are worth asking in order to contribute to building a future that we want to be part of.

In conclusion, this exploratory paper contributes to the literature on welfare technology (Peine et al, 2015; Östlund et al, 2015) by foregrounding important issues to be taken into consideration when technology is introduced and used for delivering welfare services. The paper also contributes to the literature on infrastructuring and convergence (Star et al, 2003) given that we propose that such processes need to deal with particular issues in the context of welfare services, and in particular care services. The paper focused on transparency within the organisation and further work needs both to deepen the analysis presented and to enlarge it to include transparency to the user.

Acknowledgments

This research has been made possible by funds granted to the HV3D project at Mälardalen University as well as by a Vinnova grant 2016-03170. We would also

like to thank the municipalities we are collaborating with and all the people that allowed us to interview them and observe them at work.

References

Aanestad, M. (2003). The camera as an actor: Design-in-use of telemedicine infrastructure in surgery. *Computer Supported Cooperative Work*, vol. 12, pp. 1-20.

Bijker, W. E., Hughes, T. P., and Pinch, T. J. (Eds.). (1987): *The Social Construction of Technological Systems. New directions in the Sociology and History of Techology*, MIT Press, Camridge and London.

Bowker, G. C. and Star, S. L. (2000): *Sorting things out: Classification and its consequences*. MIT press.

Bratteteig, T. and Wagner, I. (2013). Moving healthcare to the home: the work to make homecare work. *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work*, 21-25 September 2013, Cyprus.

Breskovic, I., Pinatti De Carvalho A. F., Schinkinger, S. and Tellioglu, H. (2013): 'Social awareness support for meeting informal carers' needs: Early development in TOPIC'. Paper presented at ECSCW Conference, Paphos, Cyprus, 21-25 September.

Cooren, F., Brummans, B.H. J. M. & Charrieras, D. (2008): 'The coproduction of organizational presence: A study of Medecins Sans Frontieres in action', *Human Relations,* vol.61, no.10, pp.1339–1370.

Corbin, J. and Strauss, A. (1985): 'Managing chronic illness at home: Three lines of work', *Qualitative Sociology*, vol. 8, no 3, pp. 224–247.

Cozza, M., Crevani, L., Hallin, A. and Schaeffer, J. (2016) 'Ageing and assistive technologies: An epistemological framework', paper presented at IFIP WG 8.2 conference in Dublin, Ireland, December.

Crevani, L and Cozza, M (2018): 'Complementary representational practices for articulating matters of concern', paper presented at the PIN-C conference 2018 in Eskilstuna, Sweden, January.

Czaja, S. J. and Schulz, R. (2006): 'Innovations in technology and aging', *Generations*, vol. 30, no. 2, pp. 6-8.

Czarniawska, B. (2008): 'Organizing: how to study it and how to write about it', *Qualitative Research in Organizations and Management: An International Journal*, vol. 3, no.1, pp. 4-20.

Dubois A. and Gadde L.-E. (2002): 'Systematic combining: An abductive approach to case research', *Journal of Business Research*, vol. 55, no. 7, pp. 553–560.

Gherardi, S. (2010): 'Telemedicine: A practice-based approach to technology', *Human Relations,* vol. 63, no. 4, pp. 501-524.

Hanson, E., Magnusson, L. and Sennemark, E. (2011): 'Practice concepts and policy analysis', *The Gerontologist,* vol. 51, pp. 561-570.

Hartswood, M., Procter, R., Rouncefield, M., and Slack, R. (2003): 'Making a case in medical work: implications for the electronic medical record', *Computer Supported Cooperative Work (CSCW)*, vol. 12, no. 3, pp. 241-266.

Kylberg, M., Löfqvist, C., Tomsone, S., Phillips, J., Liepina, Z., and Iwarsson, S. (2015): 'A European perspective on the service delivery systems for assistive technology – differences and similarities between Latvia and Sweden', *Journal of cross-cultural Gerontology,* vol. 30, pp. 51-67.

Latour, B. (2004): 'Why Has Critique Run out of Steam? From Matters of Fact to Matters of Concern', *Critical Inquiry,* vol. 30, no. 2, pp. 225-248.

Lave, J. and Wenger, E. (1991): *Situated learning: Legitimate peripheral participation*. Cambridge university press.

MacManus, S.A., Polk, A.L. and Bonanza, D. (2013): 'The politics of ageing', in D. Dannefer and C. Phillipson (eds.): *The Sage Handbook of Social Gerontology,* Sage, London, 2013, pp. 641-652.

Nordic Centre for Welfare and Social Issues (2010): *Focus on welfare technology*. December Retrieved from http://www.nordicwelfare.org/PageFiles/5488/Velferdsteknologi_eng.pdf

Orlikowski, W.J. and Scott, S.V. (2008): 'Challenging the separation of technology, work and organization', *The Academy of Management Annals,* vol. 2, pp. 433-474.

Östlund, B., Olander, E., Jonsson, O. and Frennert, S. (2015): 'STS-inspired design to meet the challenges of modern aging. Welfare technology as a tool to promote user driven innovations or another way to keep older users hostage?', *Technological Forecasting & Social Change*, vol. 93, pp. 82-90.

Peine, A., Faulkner, A., Jæger, B., and Moors, E. (2015): 'Science, technology and the 'grand challenge'of ageing—Understanding the socio-material constitution of later life', *Technological Forecasting & Social Change*, vol. 93, pp. 1-9.

Pink, S. and Morgan, J. (2013): 'Short-Term Ethnography: Intense Routes to Knowing', *Symbolic Interaction,* vol. 36, no. 3, pp. 351-61.

Pipek, V. and Wulf, V. (2009): 'Infrastructuring: Towards an Integrated Perspective on the Design and Use of Information Technology', *Journal of the Association of Information Systems,* vol. 10, no. 5, pp. 306-332.

Søndergård, D.C. (2014): *Future challenges and the role of welfare technology*. Nordic center for welfare and social issues. Retrieved from http://www.nordicwelfare.org/PageFiles/47/Future%20Challenges%20and%20the%20Role%20of%20Welfare%20Technology%20in%20the%20Nordic%20Countries.pdf

Star, L. and Bowker, G. (2002): 'How to infrastructure', in L. Lievrouw and S. Livingstone (eds.): *Handbook of New Media: Social Shaping and Consequences of ICTs*, Sage, London, pp. 230-245.

Star, S.L., G. Bowker and L. Neumann (2003): 'Transparency Beyond the Individual Level of Scale: Convergence Between Information Artifacts and Communities of Practice', in A.P. Bishop, B.P. Buttenfield and N.V. House (eds.): *Digital Library Use: Social Practice in Design and Evaluation*, MIT Press, Cambridge, pp. 241-269.

Star, S. L. and Ruhleder, K. (1996): ‘Steps toward an ecology of infrastructure: Design and access for large information spaces’, *Information systems research*, vol. 7, no. 1, pp. 111-134.

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

Bustamante Duarte, AM.; Degbelo, A.; Kray, C. (2018): Exploring Forced Migrants (Re)settlement & the Role of Digital Services. In: Proceedings of 16th European Conference on Computer-Supported Cooperative Work - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2018_7

Exploring Forced Migrants (Re)settlement & the Role of Digital Services

Ana Maria Bustamante Duarte, Auriol Degbelo, Christian Kray
Sitcom Lab, Institute for Geoinformatics, University of Münster, Münster, Germany
(bustamante | degbelo | c.ckray) @uni-muenster.de

Abstract. In recent years, large numbers of forced migrants have arrived in urban areas all around the world. Access to relevant information and suitable technology can help forced migrants, mainly refugees and asylum seekers, to cope with several of the challenges they face in this process. We conducted a qualitative study with ten forced migrants and six social workers and a staff member of a collective lodging for young forced migrants in Münster, Germany. The goal was to identify challenges and needs in this specific context, find criteria for assessing digital support services for forced migrants, and suggest general aspects of improvement. We analyzed 36 existing mobile applications and web services useful for forced migrants upon arrival and during (re)settlement. Our results highlight some critical issues to be addressed through digital services for forced migrants regarding information reliability, timeliness, and complexity, as well as an occasional lack of experience with geospatial services.

Introduction

Forced migration is a global phenomenon: around 65.6 million people in this situation were registered in 2016 (UNHCR, 2016). Germany, for example, registered 722.370 asylum applications for the same year (BAMF, 2017). Consequently, there is a need for a better understanding of this phenomenon and for strategies and (technology-related) tools to support forced migrants.

Forced migrants face distinct migration phases and information landscapes while going through diverse stages of awareness (Kennan et al., 2011). In this process, forced migrants show high levels of resilience and their strong collaboration and support dynamics (Fisher et al., 2013; Marlowe, 2010). Nonetheless, they also encounter several challenges such as social isolation (Almohamed, 2016; Andrade and Doolin, 2016), information poverty (Caidi et al., 2010), cultural barriers, limited proficiency in the local language (Brown and Grinter, 2016), lack of trust (Almohamed and Vyas, 2016), as well as limited access to services and health care (Talhouk et al., 2016a,b). Various technological solutions have been developed to help forced migrants tackle these challenges e.g., (Brown and Grinter, 2016; Talhouk et al., 2017; Shankar et al., 2016; Xu and Maitland, 2016).

Despite prior findings regarding challenges and needs of forced migrants, further work is needed due to the heterogeneity of the group, and the large number of factors influencing their situation. In this paper, we contribute to research with forced migrants in two ways: (i) we gathered their challenges and needs upon arrival and during (re)settlement in Münster, Germany; (ii) we used these findings to conduct an exploratory analysis of 36 services designed for, or useful to forced migrants. This analysis identified strategies used by these digital services to address the challenges from forced migrants, as well as their current gaps. Our results can benefit forced migrants and all relevant actors as they work towards solutions aimed at helping forced migrants upon arrival, and during the first stages of the (re)settlement process.

Related work

As stated previously, forced migrants go through diverse information stages during their displacement. Kennan et al. (2011) suggested three stages (*transitioning, settling in,* and *being settled*) which are defined as "cyclical and iterative". The *transitioning* phase occurs prior to forced migrants' arrival to their host country. In this stage they are "seeking" and receiving information about their host country. In the *settling-in* stage, forced migrants are "oriented" and no longer limited to the information that is being provided to them but they actively extend it in scope and sources. Finally, in the *being settled* phase forced migrants have a clearer understanding of the information landscape while "constructing an internal map" about it and sharing this information with others.

Technology, as stated by previous work, facilitate a sustainable integration of refugees in their new place. AbuJarour and Krasnova (2017) observed that technology can enable numerous capabilities for Syrian refugees in Germany: social connectivity, effective telecommunication, safety and emergency services, mobility, translation services, the participation in an information society and in educational programs, the communication with the government, crowdsourcing, as well as maintaining refugees' cultural identity. Moreover, based on their experience with the Come_IN initiative in Germany, Weibert and Wulf (2010)

concluded that computer technology is apt to promote both sustainable structural and cultural integration. The authors also suggested that a computer-based project can serve to establish and strengthen intercultural relationships in a neighborhood. Further work into the adaptation of the Come_IN approach to a refugee camp in the West Bank shows that computer clubs can promote mutual learning between children refugees and student volunteers (Yerousis et al., 2015; Aal et al., 2014). Children refugees may extend their perspective over the boundaries of the refugee camp while acquiring new skills and contacts; and student volunteers gain a better understanding of the needs and struggles of the camp inhabitants. Lastly, Bustamante Duarte et al. (2018) conducted workshops with young forced migrants and young locals in Münster, Germany in order to codesign a mobile tool to support the former upon arrival and first stages of (re)settlement. Their study pointed out that the combination of participatory design and participatory research strategies is useful for engaging and building trust with young forced migrants while designing digital services for them.

Moreover, technology can also be helpful for forced migrants during their daily life activities in their new environment. For example, Baranoff et al. (2015) proposed a mobile service, *Lantern*, which helps refugees to navigate and learn about their new environment obtaining context-specific help using Near Field Communication (NFC) technology. Brown and Grinter (2016) mentioned several benefits of *Rivrtran*, a tool to facilitate engagement between refugees and their mentors (i.e., American families) during the (re)settlement process. The tool helped forced migrants to articulate their needs better, jointly formulate goals with their mentors, and initiate communication. Schreieck et al. (2017) developed *Integreat*, a mobile app which aims to provide local information for forced migrants in several cities in Germany. In their work, they derived a series of design principles for applications which aim to transmit information to a culturally diverse audience. Also, Ngan et al. (2016) developed *Moin*, a mobile app to support informal language learning and integration of young refugees in Bremen, Germany. In a more general perspective, Harney (2013) pointed out that mobile phones offer the possibility of mobilizing personal networks and aid forced migrants reducing the fears and uncertainties they have about their new place of (re)settlement. Gifford and Wilding (2013) indicate that technology offers new possibilities of imagining social horizons beyond the constraints of their settlement context to young people with refugee backgrounds. Phillips (2013) examined the impact of remote telephone interpreting in the (re)settlement experience of refugees. The author argued that despite the widespread view that on-site interpreters are always preferable to remote interpreters, refugees may be better served by telephone interpreters. Telephone interpreting services offer two benefits for forced migrants: recognition of their individual needs by the polity, and the safe negotiation of identity through the (re)settlement process. Moreover, Hashemi et al. (2017) analyzed mobile apps for language training and information provision regarding the host society in Sweden through the Technological, Pedagogical, Linguistic and Cultural model (TPLC-model). Their results evidenced that mobile

digital services related to translation and language training are common but not apps regarding societal information for the Swedish case.

Despite the key role technology can play easing the lives of forced migrants, it can also cause further difficulties. For instance, Schmitt et al. (2016) examined the technical technology infrastructure of the Za'atari refugee camp. Some issues were identified such as the uneven spatial distribution of signal coverage and carrier congestion which affect the life of the refugees in the camp. According to Wilding and Gifford (2013), technology can make it easier for others to make demands from distance potentially straining social relationships between forced migrants, and the kin at their home country. As this section illustrated, technology plays an important role in the lives of forced migrants. Understanding challenges and needs of forced migrants (both contextual, and universal), and key gaps of services designed for them is therefore important for the development of applications which effectively support forced migrants at different stages of the migration process.

Approach and methodology

This work aims to 1) identify challenges and needs of forced migrants in Münster, Germany; 2) explore strategies to assess how digital services are addressing these; and 3) define initial aspects for improvement in these services. To achieve this, we applied a three-step method. First, we carried out a qualitative study with forced migrants and actors involved in their process of (re)settlement. Second, we used the outcome of the first study to derive criteria for an exploratory systematic analysis of existing services aimed at supporting forced migrants in Münster.

Interview study with forced migrants and other relevant actors

Context

We conducted a series of interviews between January and November 2016 with forced migrants (N=10), social workers (N=6) and a collective lodging support staff member in Münster. We wanted to identify the needs, challenges and information communication strategies of forced migrants during their initial stages of (re)settlement in Münster. In 2016, Münster registered 2412 asylum seekers. In 2017, 942 refugees were officially registered.

Participants

For this first study, we recruited participants from two groups: 1) forced migrants in Münster (four individual interviews and two on-site group interviews), and 2) social workers and lodging support staff members (four individual interviews and one on-site group interview). Forced migrants in this article refer to both refugees and asylum seekers. We used snowball sampling to recruit participants from both groups. Ten forced migrants (two females and eight males, aged between 19 and

46 years old) participated in the interviews. Eight participants were from Syria, while two were from Albania and Eritrea. All participants had completed high school education. Four participants had achieved university degrees or were pursuing university education before fleeing their country. At the time of the interview, participants had been in Münster between 7 and 17 months. The six social workers and the support staff member worked in two different types of residences for forced migrants: 1) collective short-term shelters (N=2), and 2) collective lodgings where forced migrants stay until their asylum claim response arrives (N=5).

Materials and Procedure

The interviews were semi-structured and included questions related to the forced migrants' life in Münster: their challenges, needs, means for searching and accessing information and services in the host city, along with their education and technology background. The interviews lasted between 25 and 50 minutes. All interviews were conducted in English. One of the two group interview sessions with forced migrants was assembled organically on-site. Some participants brought other forced migrants living in the lodging to attend or translate to other languages.

Analysis

We used MAXQDA for the analysis of the collected qualitative data in three iterative cycles. We particularly focused on finding patterns related to forced migrants' access and use of information for the *transitioning* and *settling-in* stages (Kennan et al., 2011). We followed a descriptive coding method (Saldaña, 2009) for the first and second iteration, which resulted in inductive (emerging) and deductive (a priori) categories (Flick et al., 2004). The deductive categories were based on the results of previous research on forced migrants and their information technologies and communication landscapes, i.e., (AbuJarour and Krasnova, 2017; Talhouk et al., 2016b; Brown and Grinter, 2016; Andrade and Doolin, 2016; Kutscher and Kreß, 2016; Baranoff et al., 2015; Lloyd et al., 2013; Kennan et al., 2011; Caidi et al., 2010; Caidi and Allard, 2005). We defined seven categories (language, information, functional literacy, technology experience, forced migrants ask for..., information communication preferences, and information sharing) for the codes, which we clustered into three main themes: *challenges, needs*, and *strategies for information sharing during the (re)settlement*. A validation of the adjusted coding scheme (after the first iteration) was carried out by all authors.

The first theme *challenges* subsumes information related to difficulties forced migrants faced when performing certain tasks during their (re)settlement process. Its general categories were classified as *a priori* and confirmed by the data obtained from the interviews. These include challenges related to language (Brown and Grinter, 2016; Andrade and Doolin, 2016; Lloyd et al., 2013; Kennan et al., 2011; Danso, 2002), functional literacy (Brown and Grinter, 2016; Kennan et al.,

2011; Lloyd et al., 2013; Caidi and Allard, 2005), information access and understanding (Lloyd et al., 2013; Caidi et al., 2010; Kennan et al., 2011; Caidi and Allard, 2005; Baranoff et al., 2015) and limited prior experience using technology (Talhouk et al., 2016b; Baranoff et al., 2015; Gillespie et al., 2016; Lloyd et al., 2013; Kennan et al., 2011). Several subcategories emerged organically from the results of the interviews such as use of geospatial services, type of information visualization, and timeliness information.

The second theme *needs* was based on the main elements (resources, strategies, tools) forced migrants in the study mentioned as relevant to them when (re)settling in Münster. These aspects consisted of two categories. First, we found *forced migrants needs* which refers to aspects they require to have a better and more effective process of arrival and (re)settlement. The codes *learning local language, accessing formal education, offline services*, and *social interaction with local community* assigned to this category were defined deductively from prior research (Andrade and Doolin, 2016; Bin Morshed et al., 2017; Kutscher and Kreß, 2016; Caidi et al., 2010; Danso, 2002; Donnelly, 2000). Additional codes such as *having translators at the beginning*, and *other services*, emerged from the interviews. Second, it is the *information communication preferences* category which aimed to identify resources forced migrants found useful for communicating information to them through technology.

The third theme *strategies for information sharing during (re)settlement* pointed to the preferred (information) communication strategies – physical or digital – that forced migrants (FMs) have during the *transitioning* and *settling-in* phases. Four categories relate to communication processes for guidance: 1) FM with FM, 2) FM with the local community, 3) FM with support staff (e.g., social workers), and 4) FM decides not to ask others for guidance. These codes emerged from the interviews.

Survey study of existing systems supporting forced migrants

Context

The main research question of this second study was "how do existing mobile services address the challenges and needs identified by forced migrants during the interviews?". Previous work (e.g., (AbuJarour and Krasnova, 2017; Schreieck et al., 2017; Andrade and Doolin, 2016; Gillespie et al., 2016; Kutscher and Kreß, 2016; Rohde et al., 2016; Xu and Maitland, 2016)) has pointed out that mobile phones, particularly smartphones, are one of the main sources for forced migrants to access, manage and communicate information. Former studies have done specific evaluations of single services created for forced migrants (e.g., *Integreat*, *Moin*). However, there has been few, if any, broader systematic assessments for this group of services considering forced migrants' challenges and needs. Our research team analyzed 25 mobile applications and 11 web platforms used by or potentially useful for forced migrants during their migration processes.

Procedure

We selected services following a two-step approach. First, we collected all services mentioned at lease twice by different participants in a list containing thirteen services as result. Second, we included services available on the web portal *Apps for Refugees* (http://www.appsforrefugees.com, accessed in March 27, 2017) which compiles services (23 mobile applications and nine web platforms) potentially useful for forced migrants. The final list contained 36 services (25 mobile applications and 11 web platforms). Since six services were dysfunctional - three web platforms (i.e., Refugeemap.com, LaGeSoNUM, and Hilfebuchen.de) and three mobile applications (i.e., helphelp2, Wülfrath hilft 2 and Hope for Austria), we only analyzed the remaining 30 services.

Analysis

We assessed the 30 services based on the results from the interviews. From it, we generated the following classification and evaluated how the selected services tackled: 1) "Limited Proficiency on Local Language", 2) "Internet Access", 3) "Information Complexity and Reliability", 4) "Prior Experience of Forced Migrants with Technology", 5) "Functional Literacy", and 6) "Strategies for Information Sharing".

Findings

Challenges and needs of forced migrants

Limited Proficiency in the Local Language

All forced migrants highlighted, as expected, language as a core challenge. Several participants related it to feelings of fear, uncertainty, and stress during their everyday interactions with the host community (e.g., doctor's appointments, grocery shopping). Based on the collected narratives, this phenomenon does not only affect their individual communication with locals but also their access to services. For example, FM_P6 stated

> "I was afraid of going to LIDL because I say, if they ask me something I don't know the language and [then] what [can I] say to them?"

Such experiences can potentially have a negative impact on forced migrants' awareness of the procedures from which they are part of. FM_P1 narrated about going to the doctor with her friend:

> "[. . .] after three times [of going there] she knows the doctors need this [..] but when they speak [to] her she cannot understand."

Additionally, having limited knowledge of the local language can be perceived by forced migrants as a loss of social status. FM_P4 stated about another resident at his collective accommodation:

> "[. . .] he was very powerful in Syria [...] but he comes [here] and he cannot [even] say "I want some food".

The assessment of digital services showed that these mainly address this challenge through:

- Education services to learn the German language: Two mobile applications promote language learning in a structured way. *Phase6 Hallo Deutsch Kinder* has as main goal language learning, while *Ankommen* aims to provide relevant but general information to guide refugees in their host country. It also provides material for German language learning.
- Tools to perform translation of phrases: Four (4/30) services offered either real-time translation (*Google Translate*) or pre-translated sentences to be used in different situations (*Refugee Phrasebook, Refugee Phrasebook -Interactive-*, and *RefuChat*). One (1/30) service connected translators with forced migrants searching for this type of assistance (*Alles Klar*).

Forced migrants also considered translation of information into their native languages as crucial during their first months. FM_P1 mentioned,

> "It is hard for [some of us the ones that] do not speak English because it is not Arabic copy for it."

In addition, the results from the survey study showed eight services (8/30) had their user interface (UI) and content in one language only (seven had only English and one only German as primary language). 21 services (21/30) had multilingual features. The languages most frequently used (which were not excluding among them) were English (26 services), Arabic (19 services), and German (17 services). 13 of these services (13/21) offered multilingual features in both UI and content. Four services (4/21) translated the content into several languages while the UI was displayed in a standard language (generally English). One service (1/21) translated only the UI components. One service (1/30) could not be fully accessed for this assessment.

Limited Internet Access

Limited internet access was a central subject for some participants (three forced migrants, and one social worker). It is a matter which also has emerged in previous research in HCI4D (Dell and Kumar, 2016) and ICT4D (Bin Morshed et al., 2017). FM_P7, for example, commented *"Without Internet, I [can't] use them"* when talking about translators and language learning services he wanted to resort to upon arrival. Limited internet connectivity can have a direct impact on the willingness of newly arrived forced migrants to explore their new host territory. In this sense, FM_P6 stated:

> "And, in camp we [didn't have] Internet there and we [were] afraid because if we [get] lost how to come back?".

The participant also added when asked about the characteristics of a useful mobile application for forced migrants,

> "should be with Internet but it [has to be] useful when we don't have Internet [too] [...]because not everybody can have Internet in their phone."

The analysis of the 30 services indicated that only thirteen supported offline use (to some extent). Seven showed some limitations when being accessed offline (see Figure 1). For example, one (*InfoAid*) required to download the daily reports (main content) beforehand for the app to operate with up-to-date information. One service (*Helping Hand*) had its content and its search services fully functional in offline mode. However, the map visualization section offered by the service was only accessible online. *Integreat* works offline but, at the time the study was conducted, few images did not load while using this mode. *Informationen für Flüchtlinge* had the option to download .pdf files while being online for later consultation in offline mode. *Ankommen* has almost all of its functionality available online, but the audiovisual resources for the language learning component had to be downloaded online prior its use in offline mode. As for the two navigation services (*osmAND* and *Google Maps*), they required downloading the area for which the map is needed prior to the offline use of the applications.

Information Complexity

Eight participants (three forced migrants/five social workers) highlighted information complexity as another significant challenge. Two factors need to be considered when presenting information to forced migrants in Münster, a) *information overload* (see (Lloyd et al., 2013; Kennan et al., 2011)), and b) the *type of information's visualization*. Concerning the first factor, SW_P4 expressed

> "I think at the beginning for the people it is so many information, and it is very tiring that are so many things"

The results from the interviews, suggest that the present issue affects mainly *compliance information* which relates to regulations, policies, and procedures of the host country (Lloyd et al., 2013). Regarding this, FM_P2 highlighted,

> "I can speak German but I find [difficult to] to translate official papers."

About the second factor, *type of information's visualization*, one clear example from the interviews focused on geospatial information. FM_P1 alluded to it, while referring to the city's buses routes maps,

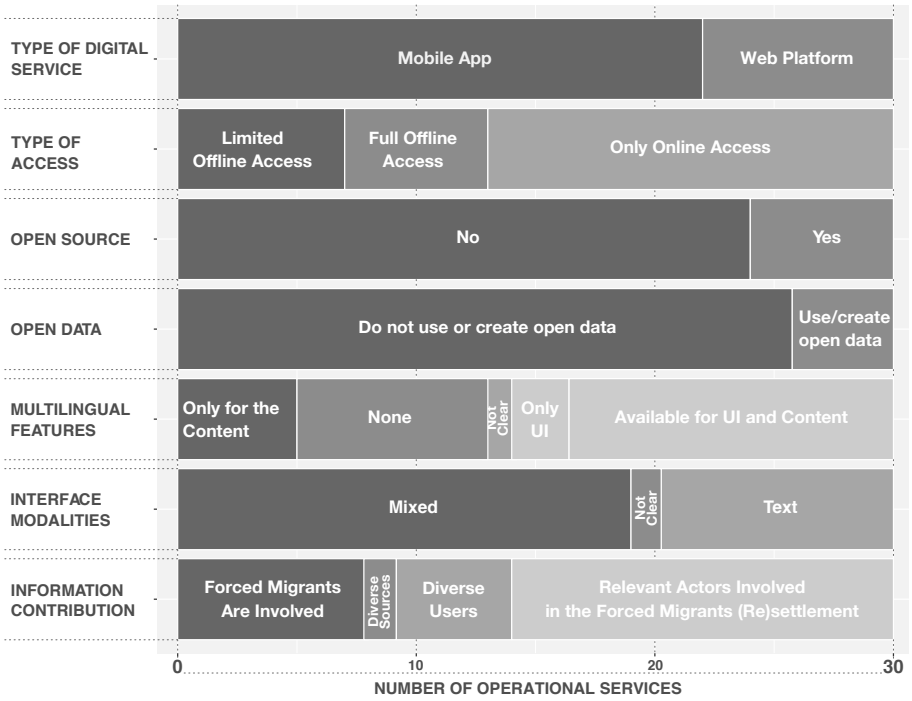

Figure 1. Overview of Main Results of the Survey Study on Operational Existing Systems.

> "I cannot understand it, because is green, and blue and pfff... it makes you [feel] terrified."

Overall, both groups of interviewees expressed that it was not always clear to forced migrants how to effectively search for places that address their needs, how to get there, and how to ensure they will actually arrive at the right place. SW_P7 indicated,

> "if you have people that is new here in Münster, they ask you everything, where is the doctor? where is the Sozialamt? [..] where can I buy that? which bus I have to take to go there?".

Regarding information communication modalities, all surveyed services used text resources (to some extent). We classified services as entirely text-based when the primary, and sometimes sole, mean for information communication was text. 10/30 assessed services fit in this category. The remaining 19 services used a variety of resources and combinations to help forced migrants overcome information complexity. From these, text (from keywords to paragraphs) was combined with: icons (3/19); audio (2/19); images and icons (5/19); audiovisual (1/19); audiovisual and icons (1/19); audiovisual, images, and icons (2/19); images, icons and audio (1/19); maps (1/19); maps and images (1/19); maps and icons (1/19), as well as maps, icons and images (1/19). Among these 19 services, seven had non-text-based resources as their primary mean for information communication using a limited amount text. One service could not be fully

analyzed in this study since the registration process did not work during our testing. Though it is unclear which of these ways of conveying information most effectively help forced migrants to deal with information complexity, our results illustrate the diversity of strategies used by existing services when communicating information to forced migrants.

Information Reliability and Timeliness

Two concerns were raised by five participants during the study regarding the trustability and timeliness of the information available. Regarding the first concern, the narratives from three forced migrants hinted at misinformation and misconception occasionally present among the forced migrants. One of them, FM_P4 said,

> "it was a common talked between the refugee[s] that [if] you tell the bus driver you are a refugee he will let you [in], that it is ok."

Another example was given by a social worker who said about forced migrants with babies that,

> "at the beginning, they need[ed] to go every week to the doctor and they didn't go to, because they thought they need[ed] to pay [...] "

The availability of up-to-date information was reported by two social workers interviewed. For instance, SW_P4 and SW_P5 stated during a group interview,

> "it would be very important that the information that are on these pages [websites] [is] updated because, for example, we got always [. . .] a list and we call[ed], and everything was already full, or it was old, so I think this can be very frustrating if you have this platform and everything is out [of] date."

We can thus conclude that frequent and timely updates of available information seem to be important to effectively address the needs of forced migrants in Münster. The analyzed services had some features towards addressing these two concerns. Regarding reliability, several services provide detailed information about topics which otherwise could be subject to inaccurate assumptions. For example, FM_P4, who was formerly quoted on the myth about the bus tickets stated,

> "The Ankommen app was telling us this is forbidden this is ok, this doesn't matter, [or] this might [bring] some problems for you. "

Also, some services provide details on the creators and contributors of the information they provide, promoting thereby the service's transparency along with the users' awareness about the information's source. The study identified seven services that clearly indicated forced migrants' involvement. In eight (8/30)

services (*Alles Klar, Welcome-Münster, Refuchat, Refoodge, Refugees-Welcome.net*, and *Refunite*) forced migrants partially contributed to their content. One more service (*Refugee Center Online*) had a combined approach where joined official inter-institutional data had a crowdsourced data curation process carried out by migrants (particularly refugees). As for *Integreat*, the creators did a survey among refugees to gathered requirements which they used to design the mobile application.

Collaborative strategies for the data creation of some services might enhance the potential for timely information. The larger the number of contributors from the group of interest or related actors, the greater the possibilities of having new and relevant information. Four services (4/30) were, distinctly, built on open data. Open data represents a way to promote collaboration between actors since it allows to jointly work upon, verify, and improve the data. From these, one (*Refugees Phrasebook* web-platform) created its own data and chose an open format through a CCO license, while three mobile services (*osmAND*, *Refugees Phrasebook-Interactive-* and InfoAid) draw upon data created in other platforms (*Open Street Map* and *Refugees Phrasebook* respectively) to build their services. The stand of the remaining 26 services regarding open data creation or use was not clear from their available documentation.

Limited Experience of Forced Migrants with Geospatial Services

Three forced migrants reported also having some difficulties using geospatial services. FM_P1 indicated

> "Google Earth [referring to Google Maps] I did not use it before. I did not need it before. In my land [country], I know everything. It is a small land; you do not get lost easily. Here you get lost easily."

Also, FM_P7 expressed when asked about how he found places in the city

> "I [search] in Google Maps, but it is not reality. [...] Maybe it is that I am not sure how to use it, so I like information from people."

Furthermore, three forced migrants mentioned learning how to use these services at diverse stages of their migration; part of them did this after fleeing from their country. FM_P2, who acted as FM_P3's interpreter during the interview, commented about FM_P3's use of Google Maps

> "[At] the beginning, he did not know how to use it, but he came with other people they helped him [...]"

Despite this difficulty, mobile geospatial services seemed to be highly useful for forced migrants. About it, FM_P6 when narrating how she and her husband move around:

> "But, just with Google Maps, it helps us a lot. [...] We don't know how to go there besides Google Maps."

While the feedback overall was positive, some also highlighted shortcomings:

> "Sometimes in Google Maps are not all the places where we want to go"

Difficulties Due to Limited Functional Literacy

Reading and writing in German as well as in the forced migrants' mother tongues are skills that become essential for their everyday lives in Münster. Functional literacy was referred to by four social workers during the interviews. SW_P3, indicated that in several collective lodgings forced migrants are receiving "Alphabetisierungkurs" where they can *"go and learn from scratch"*. Additionally, SW_P2 stated that it is of high relevance for city institutions *"make sure [everything] is in every language"* when providing information to forced migrants. However, the participant also said,

> "when people [don't] know how to read or write [in] their own language, then we have a problem again."

Few forced migrants in Münster have limited functional literacy and classes in the collective accommodation as well as in schools are organized to support them in this matter. Nonetheless, according to the narratives of the social workers interviewed, at the beginning they have difficulties accessing information and learning the local language due to it. Concerning the analyzed services, many did not directly state a goal of creating a service which could be used by users with limited functional literacy except two services (2/30) (*InfoCompassBerlin* and *Refugees Center Online*). Both have as part of their mission to convey information to larger audiences irrespective of their educational background.

Strategies for Information Sharing in the (Re)Settlement

Four main types of strategies for guidance across the new information landscape (Kennan's *transitioning* and *settling-in*) were suggested during the interviews with forced migrants: 1) from other forced migrants (FM-FM), 2) from local community and volunteers (FM-Local Community), 3) from members of official institutions (FM-Official Institutions), and 4) those who do not ask for guidance, but rely only on the information they gathered.

The first type of collaboration strategy (FM-FM) was the most commonly highlighted by participants (five forced migrants/four social workers). The similarity in the cultural and social background and the presence of a common language hinted at a solid foundation for asking first others in the same situation. For instance, FM_P1 recounted when we asked about how she found her way to supermarkets and shops in the first months *"We go sometimes by bus and we learn*

each other." Likewise, FM_P2 stated (when asked about how the exploration of the city was done during the first days),

> "the people moves in groups, not one man, or two [...]they move in groups so the groups lead the other people, Google is used in case of an emergency or something if someone is lost they can find their way back."

The second guidance strategy (FM-local community) was reported by several participants (six forced migrants/five social workers). Forced migrants valued the various efforts done by the local community and volunteers (e.g., students' initiatives) to support them and welcome them in the new community. Such interactions were strengthened through the use of online platforms and groups in social media. For example, FM_P5, P8, P9 and P10, during the group interview highlighted their use of

> "the *Welcome Münster* service a lot since people helps you there [...] Germans organize parties, we see it, and we go there."

FM_P2, acting as an interpreter during the group interview, mentioned about FM_P3's experience,

> "German people can understand English, his friend speaks a bit English [and] German people wanted to help."

Using this information sharing and collaboration strategy seems to promote social relationships between forced migrants and members of their host community. It can also guide forced migrants while navigating the new information landscape as well as the culture of the host community. As an illustration, FM_P4 commented

> "me and my brother [...] we found some nice Germans [who] spoke English, and they helped us a lot translating [...] guiding us; they helped my brother to get his flat."

Regarding the third information sharing strategy (FM-Official Institutions), four forced migrants recognized social workers in Münster as great facilitators and intermediaries to access information. FM_P6, for example, said

> "[..]social workers [...] are great, they discuss with us, they make all the things."

Moreover, participants (two forced migrants/two social workers) also recognized the help of well-established civil institutions such as *"the GGUA. [...], and AFAQ"* [1](FM_P5, P8, P9, and P10).

[1] The *Gemeinnützuge Gesellschaft zur Unterstützung Asylsuchender e.V.* (GGUA) is a registered association which supports asylum seekers, refugees and migrants in Münster by offering them social and legal advice (see http://www.ggua.de). The *Verein für Kulturelle und Gesellschaftliche Zusammenarbeit* (AFAQ e.V.) is an organization which promotes intercultural and social cooperation in Münster (see http://afaq-verein.de/?lang=en; last accessed: April 18, 2018).

Lastly, two forced migrants indicated they preferred not to ask for guidance. About it, FM_P6 stated (following a previous question where she raised concerns of asking locals)

> " I don't like it, because if I ask for something and they don't give me help, then I'll feel bad."

11 services (11/30) promoted actions which involved active communication processes between forced migrants and other relevant actors. Five of these services (5/11) actively encourage bidirectional information sharing between forced migrants and the local community (FM-Local Community). Specifically, *Refugees-Welcome.net* encourages locals to offer housing options to refugees and asylum seekers. *Refoodge*, promotes dinners between forced migrants and locals. *Welcome-Muenster.org* organizes social events (parties, city tours, sports events) for both groups. *Alles Klar* connects forced migrants with local translators in their cities of arrival. Finally, *Refuchat* has ready-sentences and a chat to ease communication between volunteers, paramedics, and forced migrants. Regarding, the *FM-FM* strategy, this is largely neglected by such services. Furthermore, the *FM-Official Institutions* strategy seems to be handled in several services through an unidirectional top-down method of information communication towards forced migrants. In those, official institutions work as providers of information for them but do not count in their services with active features for forced migrants to constantly communicate with them (e.g., chats, posts) or to adapt the services.

Discussion

Two studies were presented in this article. One identified challenges and needs of forced migrants in Münster, Germany via a series of interviews; the other one shed some light on strategies of existing platforms to address (some of) the identified challenges. The discussion in this section revolves around three aspects: new (or emerging) challenges from our data, major gaps in current services, and major implications of our results for research on collaborative technologies for forced migrants.

Emerging challenges from the interviews

Challenges related to language, functional literacy, information access and understanding, as well as limited experience using technology were mentioned in previous work and also voiced by the participants. In addition, the interviews have highlighted issues not so often documented in previous studies. In particular, they reported challenges coping with existing geospatial information, along with limited experience with geospatial services in general. This is important given that geospatial information and services are the basis for navigating, as well as

developing cognitive collages and spatial mental models of the new environment [2]. One participant referred to the diversity of colors and routes were at time confusing. One doubted that the information provided by existing geospatial services (e.g., Google Maps) reflected reality. Participants' feedback may be the result of many factors (notably information overload). Still, they remind that Shneiderman (2000)'s vision of universal usability is yet to be achieved for geospatial information.

Key gaps of current services

The majority of the services which were developed for forced migrants seemed to be focused on the Arabic speaker group which move to English or German-speaking countries. Thus, other groups of forced migrants from countries such as Somalia, Eritrea, Albania, and Afghanistan, who are also arriving in Münster are mostly not covered by these technologies. Additionally, forced migrants mentioned reduced internet availability as a core impediment for information access during the interviews. Currently, approximately one-third of all mobiles services analyzed address this issue (out of which half were having difficulties providing fully operational services in offline mode). This calls for more work producing tools supporting offline usage, to better cope with the conditions of forced migrants' life.

Implications for research on collaborative technologies for forced migrants

The analysis highlighted three key information sharing strategies: among forced migrants; between forced migrants and the local community; and between forced migrants and official institutions (via social workers as prime intermediaries). One third of the services did provide some features which can support multi-directional information exchange between the different parties. It is an interesting research question to explore how tools which support all three information sharing strategies could be designed. Further assessing the impact of these services on forced migrants life (e.g., via ethnographic studies) would be valuable for our understanding of CSCW-related systems in non-work settings.

Several forced migrants reported difficulties while using geospatial information or services. Despite not being widely supported by the studied digital services, the *FM-FM* strategy was used by forced migrants either to ask others for directions, to explore the city in groups, or teach each other how such geospatial services worked. Supporting this kind of collaboration into services and leveraging it during the design process, may lead to more effective tools that are better adapted

[2] Cognitive collages and spatial mental models were presented in (Tversky, 1993) as two metaphors describing people's knowledge about their environment: the former refers to distorted, incomplete spatial knowledge, while the latter denotes coherent mental representations of spatial relations among landmarks.

to the needs and practices of forced migrants. Developing and incorporating more flexible and collaborative visualizations of the information (e.g., (Brodlie, 2005; Isenberg et al., 2011)) and geovisualizations (e.g., (Fechner and Kray, 2014; MacEachren and Brewer, 2004; Nöllenburg, 2007)) also constitutes a promising line for future work. Similarly, adding location-based features to promote co-located information exchanges among forced migrants could help to address information complexity upon arrival by filtering information based on their current immediate context.

Moreover, the small number of services using open source data (3/30) was revealing. Since forced migration is a recurring phenomenon, the use of open data (and open source platforms) could be useful to capitalize on past experiences. At the moment, most of the services currently available do not have a clearly documented policy as regards the license of their data. Since data is copyrighted by default, re-use of this existing data can barely happen. Research on services for forced migrants may thus benefit from a more open approach towards data and service sharing.

Limitations

The limited number of participants is one of the drawbacks of this exploratory study influenced by several factors. First, forced migrants in the initial phases of (re)settlement have a variety of daily activities which are a priority (e.g., doctor's appointments, asylum claim procedures). Hence, setting up fixed interviews was a complex task. Second, we specifically targeted forced migrants who spoke English. They often act as social and communication intermediaries in the lodgings being aware of issues several forced migrants are facing. Thus, most insights obtained were from participants who spoke English. Fewer responses were gathered from participants speaking German, Arabic, and Farsi, consequently under-representing the opinion of these groups of forced migrants.

Additionally, some challenges and needs (e.g., limited functional literacy, need for updated information) were obtained only from the interviews with social workers or staff members at the lodgings. Therefore, further explorations are needed to study these aspects from the forced migrants' perspective. Furthermore, only two sources were used to gather the surveyed digital services. Additional studies could include a larger sample, and assess the services directly through usability tests with the forced migrants. Finally, our study assessed mainly single-purpose mobile apps and web platforms and did not explore more general platforms which are also useful for information access and sharing among forced migrants (e.g., Facebook, Whatsapp, or Telegram). Analyzing these platforms would have provided a more complete picture of issues, and best practices of services used by forced migrants during their resettlement.

Conclusion

Previous work has identified challenges and needs of forced migrants in countries such as Australia, Canada, New Zealand, and the United States. The work reported in this paper complements these by identifying challenges, needs and collaboration strategies of forced migrants (FM) in Münster, Germany. The participants mentioned challenges identified in other contexts such as language, functional literacy, information access and understanding. In addition, they reported issues not so often documented in previous studies such as difficulties coping with existing geospatial information and limited experience with geospatial services. Based on the above information, we assessed 25 mobile applications and 11 web platforms which can support forced migrants in their (re)settlement. Our analysis highlighted the need for exploring information visualization strategies which consider information overload. It also calls for tools which favor reliability and timeliness of the available information, and which promote the information sharing strategies of forced migrants identified (i.e., FM-FM, FM-Locals, and FM-Official Institutions). Further explorations on this matter would be valuable towards CSCW approaches for forced migrants during their (re)settlement. Finally, the analysis calls for more services that work offline, and a more open approach towards data sharing to enable the community to better capitalize on past experiences.

Acknowledgments

The authors gratefully acknowledge funding from the European Union through the GEO-C project (H2020-MSCA-ITN-2014, Grant Agreement Number 642332, http://www.geo-c.eu/).

References

Aal, K., G. Yerousis, K. Schubert, D. Hornung, O. Stickel, and V. Wulf (2014): 'Come in Palestine: Adapting a German Computer Club Concept to a Palestinian Refugee Camp'. In: *Proceedings of the 5th ACM International Conference on Collaboration Across Boundaries: Culture, Distance and Technology*. pp. 111–120, ACM.

AbuJarour, S. and H. Krasnova (2017): 'Understanding the role of ICTs in promoting social inclusion: The case of syrian refugees in Germany'. In: *Twenty-Fifth European Conference on Information Systems (ECIS)*. Guimarães, Portugal.

Almohamed, A. (2016): 'Designing for the Marginalized : A step towards understanding the lives of refugees and asylum seekers'. In: *DIS*. pp. 165–168.

Almohamed, A. and D. Vyas (2016): 'Vulnerability of displacement: challenges for integrating refugees and asylum seekers in host communities'. In: H. B. L. Duh, C. Lueg, M. Billinghurst, and W. Huang (eds.): *Proceedings of the 28th Australian Conference on Computer-Human Interaction - OzCHI '16*. Launceston, Tasmania, Australia, pp. 125–134, ACM Press.

Andrade, A. D. and B. Doolin (2016): 'Information and Communication Technology and the Social Inclusion of Refugees'. *MIS Quarterly*, vol. 40, no. 2, pp. 405–416.

Arnsberg, B. 'Weitere Zahlen -FlüAG -Verteilstatistik-'.

BAMF (2017): 'Das Bundesamt in Zahlen 2016 - Asyl'. Technical report, Bundesamt für Migration und Flüchtlinge.

Baranoff, J., R. I. Gonzales, J. Liu, H. Yang, and J. Zheng (2015): 'Lantern: Empowering Refugees Through Community-Generated Guidance Using Near Field Communication'. In: *Extended Abstracts of the ACM CHI'15 Conference on Human Factors in Computing Systems*, Vol. 2. pp. 7–12.

Bin Morshed, M., M. Dye, S. I. Ahmed, and N. Kumar (2017): 'When the Internet Goes Down in Bangladesh'. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17*. Portland, OR, USA, pp. 1591–1604, ACM.

Brodlie, K. (2005): 'Models of collaborative visualization'. In: J. Dykes, A. M. MacEachren, and M. J. Kraak (eds.): *Exploring Geovisualization*. Elsevier, Chapt. 23, p. 463.

Brown, D. and R. E. Grinter (2016): 'Designing for Transient Use: A Human-in-the-loop Translation Platform for Refugees'. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. pp. 321–330, ACM.

Bustamante Duarte, A. M., N. Brendel, A. Degbelo, and C. Kray (2018): 'Participatory Design and Participatory Research'. *ACM Transactions on Computer-Human Interaction*, vol. 25, no. 1, pp. 1–39.

Caidi, N. and D. Allard (2005): 'Social inclusion of newcomers to Canada: An information problem?'. *Library and Information Science Research*, vol. 27, no. 3, pp. 302–324.

Caidi, N., D. Allard, and L. Quirke (2010): 'Information Practices of Immigrants'. *Annual Review of Information Science and Technology*, vol. 44, no. 1, pp. 491–531.

Danso, R. (2002): 'From 'there' to 'here': An investigation of the initial settlement experiences of Ethiopian and Somali refugees in Toronto'. *GeoJournal*, vol. 56, no. 1, pp. 3–14.

Dell, N. and N. Kumar (2016): 'The Ins and Outs of HCI for Development'. pp. 2220–2232.

Donnelly, A. (2000): 'Towards an Inclusive Information Society : Some Principles from the Margins'. In: *6th ERCIMWorkshop "Uster Interfaces for All"*. Florence, Italy.

Fechner, T. and C. Kray (2014): 'Georeferenced Open Data and Augmented Interactive Geo-Visualizations as Catalysts for Citizen Engagement'. *JeDEM - eJournal of eDemocracy and Open Government*, vol. 6, no. 1, pp. 14–35.

Fisher, K. E., A. P. Bishop, and P. Fawcett (2013): 'InfoMe @ Teen Design Days : A Multi-Disciplinary , Design Thinking Approach to Community Development'. pp. 25–28.

Flick, U., E. von Kardoff, I. Steinke, and B. Jenner (2004): *A Companion to Qualitative Research*, A Companion to Qualitative Research. SAGE Publications.

Gifford, S. M. and R. Wilding (2013): 'Digital escapes? ICTs, Settlement and belonging among Karen Youth in Melbourne, Australia'. *Journal of Refugee Studies*, vol. 26, no. 4, pp. 558–575.

Gillespie, M., L. Ampofo, M. Cheesman, B. Faith, E. Iliadou, A. Issa, S. Osseiran, and D. Skleparis (2016): 'Mapping Refugee Media Journeys Smartphones and Social Media Networks Research Report'. Technical Report Research Report, The Open University / France Medias Monde.

Harney, N. (2013): 'Precarity, affect and problem solving with mobile phones by asylum seekers, refugees and migrants in Naples, Italy'. *Journal of Refugee Studies*, vol. 26, no. 4, pp. 541–557.

Hashemi, S. S., N. B. Lindström, L. Bartram, and L. Bradley (2017): 'Investigating Mobile Technology Resources for Integration'. *Proceedings of the 16th World Conference on Mobile and Contextual Learning - mLearn 2017*, pp. 1–8.

Isenberg, P., N. Elmqvist, J. Scholtz, D. Cernea, K.-L. Ma, and H. Hagen (2011): 'Collaborative visualization: Definition, challenges, and research agenda'. *Information Visualization*, vol. 10, no. 4, pp. 310–326.

Kennan, M. A., A. Lloyd, A. Qayyum, and K. Thompson (2011): 'Settling in: The relationship between information and social inclusion'. *Australian Academic & Research Libraries*, vol. 42, no. 3, pp. 191–211.

Kutscher, N. and L.-M. Kreß (2016): '"Internet is the same like food" – An empirical study on the use of digital media by unaccompanied minor refugees in Germany'. *Transnational Social Review*, vol. 1674, no. June, pp. 1–4.

Lloyd, A., M. A. Kennan, K. M. Thompson, A. Qayyum, and D. Bawden (2013): 'Connecting with new information landscapes: information literacy practices of refugees'. *Journal of Documentation*, vol. 69, no. 1, pp. 121–144.

MacEachren, A. M. and I. Brewer (2004): 'Developing a conceptual framework for visually-enabled geocollaboration'. *International Journal of Geographical Information Science*, vol. 18, no. 1, pp. 1–34.

Marlowe, J. M. (2010): 'Beyond the discourse of trauma: Shifting the focus on Sudanese refugees'. *Journal of Refugee Studies*, vol. 23, no. 2, pp. 183–198.

Ngan, H. Y., A. Lifanova, J. Jarke, and J. Broer (2016): 'Adaptive and Adaptable Learning'. vol. 9891, pp. 521–524.

Nöllenburg, M. (2007): *Geographic visualization*, pp. 257–294. Berlin, Heidelberg: Springer Berlin Heidelberg.

Phillips, C. (2013): 'Remote telephone interpretation in medical consultations with refugees: meta-communications about care, survival and selfhood'. *Journal of Refugee Studies*, vol. 26, no. 4, pp. 505–523.

Rohde, M., K. Aal, K. Misaki, D. Randall, A. Weibert, and V. Wulf (2016): 'Out of Syria: Mobile Media in Use at the Time of Civil War'. *International Journal of Human-Computer Interaction*, vol. 32, no. 7, pp. 515–531.

Saldaña, J. (2009): *The Coding Manual for Qualitative Researchers*. London Thousand Oaks, Calif: Sage Ltd.

Schmitt, P., D. Iland, E. Belding, B. Tomaszewski, Y. Xu, and C. Maitland (2016): 'Community-level access divides: A refugee camp case study'. In: K. Toyama (ed.): *Proceedings of the Eighth International Conference on Information and Communication Technologies and Development - ICTD '16*. Ann Arbor, Michigan, USA, pp. 1–11, ACM Press.

Schreieck, M., J. Zitzelsberger, S. Siepe, M. Wiesche, and H. Krcmar (2017): 'Supporting refugees in every Day life–intercultural design evaluation of an application for local information'. In: *21st Pacific Asia Conference on Information Systems (PACIS)*. Langkawi, Malaysia, pp. 1–13, Association for Information Systems (AIS).

Shankar, S., H. L. O'Brien, E. How, Y. W. Lu, M. Mabi, and C. Rose (2016): 'The role of information in the settlement experiences of refugee students'. *Proceedings of the Association for Information Science and Technology*, vol. 53, no. 1, pp. 1–6.

Shneiderman, B. (2000): 'Universal usability'. *Communications of the ACM*, vol. 43, no. 5, pp. 84–91.

Talhouk, R., S. Ishtiaque Ahmend, V. Wulf, C. Crivallero, V. Vlachokyriakos, and P. Olivier (2016a): 'Refugees and HCI SIG : The Role of HCI in Responding t o the Refugee Crisis'. *CHI Extended Abstracts on Human Factors in Computing Systems*, pp. 1073–1076.

Talhouk, R., S. Mesmar, A. Thieme, M. Balaam, P. Olivier, C. Akik, and H. Ghattas (2016b): 'Syrian Refugees and Digital Health in Lebanon'. pp. 331–342.

Talhouk, R., K. Montague, and P. Olivier (2017): 'Implications of Synchronous IVR Radio on Syrian Refugee Health and Community Dynamics *'. Vol. 2017. pp. 1–10.

Tversky, B. (1993): 'Cognitive maps, cognitive collages, and spatial mental models'. In: A. U. Frank and I. Campari (eds.): *Spatial Information Theory A Theoretical Basis for GIS*. Marciana Marina, Elba Island, Italy: Springer Berlin Heidelberg, pp. 14–24.

UNHCR (2016): 'Global Report'. Technical report, United Nations High Commissioner for Refugees.

Weibert, A. and V. Wulf (2010): '"All of a sudden we had this dialogue...": intercultural computer clubs' contribution to sustainable integration'. In: P. Hinds, A.-M. Søderberg, and R. Vatrapu (eds.): *Proceedings of the 3rd international conference on Intercultural collaboration - ICIC '10*. Copenhagen, Denmark, pp. 93–102, ACM Press.

Wilding, R. and S. M. Gifford (2013): 'Introduction to the special issue on forced displacement, refugees and ICTs: transformations of place, power and social ties'. *Journal of Refugee Studies*, vol. 26, no. 4, pp. 495–504.

Xu, Y. and C. Maitland (2016): 'Communication behaviors when displaced: A case study of Za'atari syrian refugee camp'. In: K. Toyama (ed.): *Proceedings of the Eighth International Conference on Information and Communication Technologies and Development - ICTD '16*. Ann Arbor, Michigan, USA, pp. 1–4, ACM Press.

Yerousis, G., K. Aal, T. von Rekowski, D. W. Randall, M. Rohde, and V. Wulf (2015): 'Computer-enabled project spaces: connecting with Palestinian refugees across camp boundaries'. In: B. Begole, J. Kim, K. Inkpen, and W. Woo (eds.): *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. Seoul, Korea, pp. 3749–3758, ACM Press.

Jason Carter, Mauro Carlos Pichiliani, Prasun Dewan (2018): Exploring the Impact of Video on Inferred Difficulty Awareness. 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/ecscw2018- 9

Exploring the Impact of Video on Inferred Difficulty Awareness

Jason Carter, Mauro Carlos Pichiliani, Prasun Dewan
Cisco Systems-RTP, IBM Research, Univ. North Carolina-Chapel Hill
jasoncartercs@gmail.com, mpichi@br.ibm.com, dewan@cs.unc.edu

Abstract. An important issue in many forms of collaboration technology is how video can help the technology better meet its goals. This paper explores this question for difficulty awareness, which is motivated by academic and industrial collaboration scenarios in which unsolicited help is offered to programmers in difficulty. We performed experiments to determine how well difficulty can be automatically inferred by mining the interaction log and/or videos of programmers. Our observations show that: (a) it is more effective to mine the videos to detect programmer postures rather than facial features; (b) posture-mining benefits from an individual model (training data for a developer is used only for that developer), while in contrast, log-mining benefits from a group model (data of all users are used for each user); (b) posture-mining alone (using an individual model) does not detect difficulties of "calm" programmers, who do not change postures when they are in difficulty; (c) log-mining alone (using a group model) does not detect difficulties of programmers who pause interaction when they are either in difficulty or taking a break; (d) overall, log-mining alone is more effective than posture-mining, alone; (e) both forms of mining have high false negative rates; and (g) multimedia/multimodal detection that mines postures and logs using a group model gives both low false positive and negatives. These results imply that (a) when collaborators can be seen, either directly or through a video, posture changes, though idiosyncratic, are important cues for inferring difficulty; (b) automatically inferred difficulty, using both interaction-logs and postures, when possible and available, is an even more reliable indication of difficulty; (c) video can play an important role in providing unsolicited help in both face-to-face and distributed collaboration; and (d) controlled public environments such as labs and war-rooms should be equipped with cameras that support posture mining.

Introduction

Collaboration technology can directly support some form of collaboration or provide awareness (Gutwin and Greenberg 2002) of collaborators to trigger some unplanned or opportunistic form of collaboration.

Awareness technology, like technology supporting direct collaboration, has supported sharing of collaborator state captured by multiple media. For example, workspace awareness such as scrollbar awareness (Gutwin and Greenberg 2002) allows sharing of the state of collaborators' user-interface while video awareness such as video walls (Abel 1990) and media spaces (Harrison and Minneman 1990, Mantei, Backer et al. 1991) supports sharing of their physical characteristics such as posture, expression, and gaze.

Moreover, awareness technology, like technology supporting direct collaboration, can attempt to give users the feeling of "being there" in one location or go "beyond being there" (Hollan and Stornetta 1992) by supporting forms of sharing not directly provided by face to face interaction. Sharing of collaborator screens (Tee, Greenberg et al. 2006) or videos supports "being there" while sharing of read/write shadows (Junuzovic, Dewan et al. 2007), radar views/multiuser scrollbars (Gutwin and Greenberg 2002), and video silhouettes (Hudson and Smith 1996) supports "beyond being there."

"Beyond being" there awareness technology supports sharing of state computed by the software. Dewan (2016) classifies such computed awareness into derived and inferred awareness. The former presents information that is logically derived from the information about the collaborators and their activities while the latter uses data mining/machine-intelligence techniques to make inferences, which may have false positives and/or false negatives.

One form of inferred awareness supported so far is whether a remote user is facing difficulty. An important reason for awareness is to determine if collaborators are in difficulty, and offer assistance if necessary (Gutwin and Greenberg 2002). Making the computer infer difficulty can reduce the amount of information to be transmitted to a remote helper and/or relieve the helper from manually determining if collaborators are in difficulty, thereby allowing the helper to discover and process difficulties of a larger number of users.

An important issue in many forms of collaboration technology is whether and how video can help the technology better meet its goals (Tang and Minneman 1990, Tang and Minneman 1991, Fish, Kraut et al. 1992, Tang and Isaacs 1992, Isaacs and Tang 1993). This paper explores this question for inferred difficulty awareness.

Driving Problem

Our initial motivation for this work comes from research by Herbsleb, Mockus et al. (2000) and Teasley, Covi et al. (2000) that shows that as distance among programmers increases, there are fewer opportunities to offer help to team members and the productivity of programmers decreases. For example, Teasley et al indicate that if someone was having difficulty with some aspect of code, another developer in the war-room "walking by [and] seeing the activity over their shoulders, would stop to provide help."

We can attempt to make help-giving independent of distance by providing collaboration awareness - (peripheral) knowledge of the remote collaborators and their activities – to an interested observer. Previous work in this area has required the observer to manually infer the status of programmers from their videos and workspace activity (Hegde and Dewan 2008) (Guo 2015). This approach has the disadvantage that team members must allocate precious screen space to awareness information and poll for difficulties. Having difficulty, by definition, is a rare event, if programmers are given problems they have the skills to solve. Thus, polling collaborators state to deduce this status can lead to wasted effort in trying to find "needles in haystacks." Moreover, such polling may, in fact, lead to overzealous help by tutors dedicated to the task of helping (Guo 2015).

One approach to address this problem is to ask programmers to manually indicate their status. Intuition tells us that those who are willing to manually change their status are likely to not set it back, just as people forget to change their busy status in an IM tool, or turn off the "call steward" light in a plane. Prior research has, in fact, shown that manually setting a "not interruptible" flag is unreliable (Milewski and Smith 2000). In the case of the difficulty status, previous work shows that people are often hesitant about explicitly asking for help: Begel and Simon (2008) found that students and new hires are late to use help; LaToza, Venolia et al. (2006) established that programmers often exhaust other forms of help before contacting a teammate; and Herbsleb and Grinter (1999) discovered that employees are less comfortable asking remote rather than co-located teammates for help. The first author's dissertation shows that the ability to press a help button during scheduled help sessions can lead to the opposite effect in which students over-ask for help.

For these reasons, we have iteratively developed an Eclipse extension, called Eclipse Helper, that automatically detects programming difficulty, communicates this information to interested observers, and allows the observers to offer help. As reported in (Carter and Dewan 2015), the system has been used in a field study in which distributed students were offered help with their homework in response to automatically detected difficulties. There were some instances where Eclipse Helper did not predict that students had difficulty, but the next day students came to office hours for help. However, each time the tool predicted a student was in

difficulty and the student was asked if they needed help, the student answered in the affirmative. This result is consistent with previous lab studies of the tool mentioned in the first author's dissertation, which also showed the lack of false positives but the presence of false negatives. This result provides the motivation for pursuing multimodal difficulty detection.

There has been much research recently in "affect detection" (Calvo and D'Mello 2010, Graziotin, Wang et al. 2014), which makes inferences about the mental state of users interacting with computers. The state may be an event-triggered ephemeral emotion or a more persistent mood.

As far as we know, ours is the only research on affect detection (D'Mello and Kory 2012) with the driving problem of providing help to (possibly remote) programmers. It is the dual of the problem in several other efforts of determining the emotions invoked when interacting with a computer-based tutor (D'Mello and Graesser 2009, D'Mello and Kory 2012, Grafsgaard, Wiggins et al. 2014, Vail, Grafsgaard et al. 2014). Our work detects an emotion required to determine the need for intermittent tutoring, while its dual detects emotions required to determine how well continuous tutoring is working. As a result, we cannot mine interaction with the tutor – which Grafsgaard, Wiggins et al. (2014) found was a primary feature in predicting emotions.

Inferred difficulty awareness can be used instead of in addition to physical or computer-supported awareness of collaborator's activities in both industrial and educational environments. Moreover, the potential helpers can be working exclusively on the task of providing help or interrupt their task to provide help based on difficulty cues from inferred and other awareness. In the rest of the paper, we first describe how mining video changes the algorithms and results of difficulty detection, and then discuss how these results can inform the various ways in which inferred difficulty awareness can be used.

Interaction Logs

To determine the additive value of mining video, we need to choose the best baseline algorithm that does not use video. Since our field study reported above, we have changed our algorithm for mining interaction logs and verified, through the lab study described below, that it reduces the false negative rate greatly (by 53%) while slightly increasing the false positive rate (by 3%). This improved algorithm, described below, serves as the baseline for the unimodal log-based algorithm used in our comparisons.

The algorithm divides the raw interaction log into 50-event segments, calculates, for each segment, the ratios of the occurrences of certain categories of commands in that segment to the total number of commands in the segment, and uses these ratios as features over which patterns are identified. Table 1 shows the

five command categories we use in our log-based algorithm and the commands in each category. The focus-in/-out commands indicate leaving/entering a programming environment window, and the show view command indicates opening a new kind of view such as a GIT view. The names of other commands explain their functions.

Table I. Mined Commands and Command Categories.

Command Category	Commands
Focus	Focus In, Focus Out
Delete	Delete Char/String, Cut
Insert	Insert Char/String, Copy, Paste, Replace, Move Caret, Select (Text/All), Line Start
Debug	Run, Add Breakpoint, Remove Breakpoint, Hit Breakpoint, Step Into/Over, Step Return, Compilation Error
Navigation	Open File, Find, Show View

These feature were chosen based on top-down thinking and analyses of logs from our studies, which showed that in the difficulty phases we observed, insertion ratios went down and one or more of the other ratios went up. Our log-based mechanism feeds these features (ratios) to the decision tree algorithm implemented in Weka (Witten and Frank 1999) to get raw predictions. Assuming that a status does not change instantaneously, the mechanism aggregates the raw predictions for adjacent segments to create the final prediction, reporting the dominant status in the last five segments. In addition, it makes no predictions from the first 100 events to ignore the extra compilation and focus events in the startup phase and account for the fact that in this phase programmers' interaction behavior is different as they "warm up."

Our previous work on log-based difficulty detection (Carter and Dewan 2010) shows that a group model (data of all users are used for each user) worked better than individual models (training data for a developer is used only for that developer), perhaps because there is uniformity among logs of different programmers and the group model offered more training data. Therefore, this algorithm built a group model.

Other research in affect detection has also mined interaction with the computer (Grafsgaard, Wiggins et al. 2014, Vail, Grafsgaard et al. 2014). However, as far as we know, ours is the only one that segments the log based on the number of events rather than time – that is, it computes features every C commands rather than every T time units. It is not clear a time-based approach can distinguish

between idle and difficult periods. There has been some research on automatically detecting off-task work (Baker 2007, Cetintas, Si et al. 2009) but it is targeted at math tutoring systems and makes use of tutor-specific features such as the average time taken by previous students to solve each tutor-specific problem. The set of user events we mine is also very different and far bigger than the ones used by other log-based schemes. For example, Grafsgaard, Wiggins et al. (2014) mine events implying successful compilations, start of a coding activity, and end of a coding activity to detect engagement and frustration – a much smaller and different set from our event-set.

Sensors for Physical Characteristics

Previous work in affect detection has shown that certain physical characteristics mined from videos of users correlate with users' affective states. Therefore, it was attractive to explore algorithms for difficulty detection that mine promising physical characteristics. To pursue this direction, we had to determine what kind of cameras we would use to identify such physical characteristics. Our answer was guided by their availability and apparent promise. We used the Microsoft Kinect camera, a commodity item today. We also used the Creative® Interactive Gesture camera, a brand new product the second author won as an Intel award, which is also intended as a commodity product. The Kinect camera is promising because it has been used by earlier work (Grafsgaard, Wiggins et al. 2014, Vail, Grafsgaard et al. 2014) to predict various emotions, and the Intel camera is promising because it directly captures positions of various facial features and research has shown that facial features correlate with various emotions (Machardy, Syharath et al. 2012, Grafsgaard, Wiggins et al. 2014, Vail, Grafsgaard et al. 2014).

Creative® Interactive Gesture Camera

As detailed in the first author's thesis, the Intel camera gave us much worse results than the Kinect camera – so we do not discuss the details of this experiment, focusing only on explaining what went wrong. There were times, shown in Figure 1, where the camera did not capture the facial positions correctly. In the left picture, the participant leans in very close to the camera. In the middle picture, the participant puts his hand on his mouth. In the right picture, the camera recognizes the participant's hand as her face. This problem is also identified by Machardy, Syharath et al. (2012) in research on determining if a video observer is paying attention. They found that glasses and hair falling on the face reduced the accuracy of their results. The problem was less severe in their study as the subjects were passively watching the video. As we see in Figure

1, our active code composition task aggravates the problem of the facial features being obscured.

This discussion provides insight about some of the choices made in previous work regarding the kind of data mined. Two studies in which the subjects were relatively passive – MacHardy et al on video watching and Fritz, Begel et al on code comprehension - mined only physical characteristics to make their inferences. They could not mine interaction logs as these were either not available (in the video study) or had only a small number of navigation events (the code comprehension study). Pure physical features worked well because the users were relatively still.

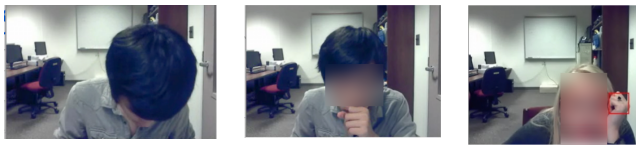

Figure 1. Some uncaptured facial positions

Kinect Camera

D'Mello and Graesser (2009) found that when participants are confused or frustrated, they tend to lean forward, and when they are bored, they tend to lean back. Based on this work, our expectation was that the Microsoft Kinect camera would measure body lean, which in turn, would correlate with difficulty. Figure 2 shows the placement of the Microsoft Kinect camera in our experiment described later. This camera captures the x, y, and z coordinates of 20 human joints at 5 fps. It could not capture all of the joints because, in our experiments, the desk that developers sat at occluded their lower body. This limitation made it difficult for the camera to capture any of the lower body joints. To support users who are sitting at a desk, it has a seated mode, which is designed to track users who sit, and does not attempt to capture joints below the hip. We used the seated mode to capture participants' joints.

Figure 2. Kinect Experimental Setup

There were times where the Kinect camera did not capture the arm and shoulder joints well, but it did capture the head joint correctly the majority of the time. Therefore, like some other studies (Grafsgaard, Wiggins et al. 2014, Vail, Grafsgaard et al. 2014), we used the distance from the head joint to the Kinect camera to measure body lean. This distance was the z coordinate of the head joint in meters.

Once we had this measure, the next question was at what rate it should be sampled. As mentioned earlier, the log-based algorithm calculates ratios every 50 events. However, in this case, this sampling scheme did not give us as good results as sampling the data every minute. Therefore, consistent with other work in posture detection (Grafsgaard, Wiggins et al. 2014, Vail, Grafsgaard et al. 2014), we used the time-based approach.

To convert the z coordinate to body lean (leaning forward, normal, leaning back) we averaged each participants' distance measure per minute and clustered each participants' data individually using the K-means clustering algorithm. Given the distance values and k, the number of clusters to produce, the algorithm partitions values into clusters based on the average Euclidean distance between them. We examined the algorithms' output with two, three, and four clusters. We used the output with three clusters because a) as mentioned above, previous work has shown that leaning back, normal, and leaning forward correlated with affective states, b) the standard deviation of one cluster was large when using two clusters, and c) four clusters offered no improvement over three. We then manually looked at representative images from each of the three clusters to see if the clusters represented the three different leans. Our samples indicated this was largely the case. Figure 3 shows examples of each type of lean from three participants captured by the Kinect camera, which were separated into three different clusters. Thus, our feature now was the cluster in which the z distance fell, which we refer to as the body lean or posture. It had three values: leaning back, normal, and leaning forward.

The next task was to choose an appropriate machine learning algorithm to mine this feature. As in the log-based scheme, we used the decision tree algorithm. Before we could feed the lean information into a machine learning algorithm, we had to determine whether we would use an individual or group model. Our intuition was to use an individual model because all developers may not be in the same posture when they are having difficulty – they may lean in different directions. However, we chose to try both individual and group models to evaluate which model would perform better

When using an individual model, we simply fed each user's per-minute lean to the decision tree algorithm. The process was more complicated in the case of the group model. Assuming that some users would lean forward when in difficulty and some would lean backward, we labeled all postures that were leaning back or

leaning forward as “NOT NORMAL,” and all normal leans as “NORMAL.” We fed this label to the decision tree algorithm.

To the best of our knowledge, the idea of using K-means clustering to classify leans and combining leaning back and forward into one feature in a group model is unique to our work.

Multimodal Algorithm

The basic idea behind combining the two unimodal algorithms is straightforward: use the features from both approaches together. However, there were two incompatibilities between them that had to be resolved. First, as mentioned earlier, in the log case, our previous work found that the group model worked better than the individual model, while, as we will see later, in the Kinect body-lean case, the individual model worked better. Second, as also mentioned earlier, in the two approaches, the features were computed at different moments - every 50 events for interaction logs and every minute for body lean.

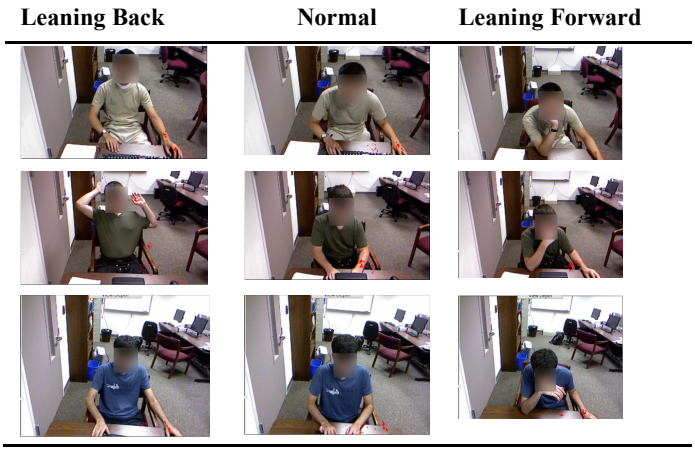

Figure 3. Examples of participants’ leaning back, normal, and leaning forward postures (Kinect camera).

After experimenting with several alternatives, we found that conforming the Kinect body-lean approach to the log-based approach worked the best. Thus, in the combined approach, we built a group model and computed Kinect features every time we had new log features. From a practical point of view, a group-model is preferable to an individual model as it does not require each new user to train the system. Other multimodal affect detection research uses time-based segmentation for all data, implicitly assuming no off-task activity.

Lab Study

To compare the unimodal algorithms with each other and with the multimodal algorithm, we needed to conduct a controlled lab study in which output of our sensors and the interaction logs were recorded. We had to choose our tasks carefully because having difficulty is a rare event. Therefore, we had to ensure developers face difficulty in the small amount of time available for a lab study, and yet did not find the problems impossible. After piloting, we settled on a problem that required participants to use the AWT/SWT toolkit to implement a GUI. Table II shows the tasks that had to be implemented in the study.

Sixteen student programmers participated in the study. They were given at least an hour and a half to complete as many subtasks as possible and were free to use the Internet. We used our previous implementation of the log-based difficulty detection mechanism (used in the field study) to record participants' programming activities and predict whether participants were having difficulty or making progress, and provided a user-interface to correct these predictions and possibly ask for help (Figure 4). The programmers could correct a wrong (a) difficulty prediction by pressing the "I am making progress" button, and (b) progress prediction by pressing the "I am solving the problem on my own" or "I am asking someone for help" buttons. If they needed help, they were instructed to discuss their issue with the first author. Help was given in the form of URLs to API documentation or code examples. By measuring how often the developers corrected their status and explicitly asked for help, we could, measure the accuracy of an inference algorithm with respect to the immediate perceptions of the developers. The videos around the difficulty points were then shown to both the participants and two observers, who confirmed them.

Table II. Mined Commands and Command Categories.

Tasks
Create a program that visually represents a car with a red body and two black tires.
Allow the user to use arrow keys to move around the car in any direction (up, forward, left, and right) by 10 pixel decrements/increments.
Allow the user to make the car a bus by clicking anywhere on the screen. A bus has an extra body that should be colored black. It should be positioned directly on top of the car. When the extra body is on top of the car, it should move with the rest of it.
Allow the user to make the bus a car by pressing the 'r' key. The extra body should be removed
Allow the users to scale up the car/bus 2X, each time they press the 'm' key
Allow the user to scale down the car/bus 2X, each time they press the 's' key

Draw a transparent square (not a rectangle) with yellow borders. The car/bus should be inside the square.
Do not allow the car/bus to go outside of the square (when moving and resizing the vehicle).

The participants' actions were captured using the Intel and Kinect cameras. However, poor lighting forced us to discard camera recordings for six of these participants. Therefore, we report the results of only ten subjects in our comparisons. The length of the study (on average 129 minutes interaction) was much larger than the duration of some other studies (e.g. 32 minutes in (D'Mello and Graesser 2009)) with more participants, and thus compensated to some extent for the discarded data. The number of used participants is typical for studying software developers Overall, the ten participants yielded 45 hours of screen and camera recordings, parts of which were examined to identify features and gather ground truth

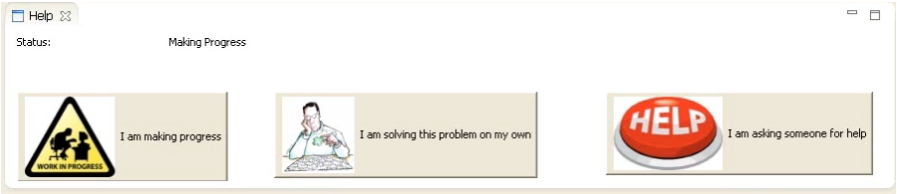

Figure 4: Correcting Predictions

Other schemes for gathering ground truth regarding affective states have relied on post-task participant surveys (Grafsgaard, Wiggins et al. 2014, Vail, Grafsgaard et al. 2014), random interruptions (Fogarty, Hudson et al. 2005), and constant observations (D'Mello and Graesser 2009). The first approach would not support our goal of making instantaneous predictions; the second would not work because, as discussed below, difficulty is a rare event and thus would not be captured by random queries; and the third approach is more labor-intensive than the one we used.

Metrics and Filtering

There were 814 predictions made for these ten users, 55 of which were difficulty predictions. Thus, the number of difficulty segments was much smaller than the number of progress segments, which is consistent with the intuition that having difficulty should be a rare event. This means that an algorithm that always predicted making progress would have very high accuracy (93% in this experiment) but not serve our goal of providing help to needy programmers. For this reason, unlike other multimodal effect detection schemes (D'Mello and Kory 2012), we calculated true/false positive/negative rates, which are presented below using confusion matrices. In the training sets, we used the Weka SMOTE filter (Witten and Frank 1999) to boost the members of the minority class (difficulty).

In our evaluations, we used the standard technique of k-fold cross-validation, which executes k trials of model construction, and splits the logged data so that 90% of the data are used to train the algorithm and 10% of the data are used to test it. We used the value 10 for k, which is typical for such validations.

Unimodal Log-Based Results

Table III shows the results (for the ten participants for which we had good camera data) with the group-based unimodal log-based algorithm. As we see here, the algorithm missed 27% of the time that participants had difficulty. Consistent with the results of our previous log-based algorithm, there were very few (3%) false positives. We extended this analysis by considering all sixteen subjects for whom we had log data. This time, the scheme gave 21% false negatives and 10% false positives. Thus, the larger data set improved the false negative rate at the cost of a higher false positive rate.

Table III: Confusion Matrix for Log Mining Algorithm.

	Predicted difficulty	Predicted Progress
Actual Difficulty	40 (True positives)	15 (False negatives)
Actual Progress	20 (False positives)	739 (True negatives)

Unimodal Lean-Based Results

The confusion matrix in Table V shows the results of the lean-based approach when we used the Kinect camera and built an individual model. As we used a time-based (rather than command-based) segmentation of the interaction sessions, the number of predictions is different from the ones in the log-based approach.

Table V: Confusion Matrix for Kinect (Individual Model)

	Predicted Difficulty	Predicted Progress
Actual Difficulty	333 (True positives)	126 (False negatives)
Actual Progress	108 (False positives)	333 (True negatives)

This scheme has the same false negative rate (27%) as the log-based approach (27%). However its false positive rate is much higher – 25% instead of 3%.

Table VII shows the results of the group model. Both the false positive (45%) and false negative (31%) rates are higher in this case, which is consistent with our intuition about individual lean idiosyncrasies.

Table VI: Confusion Matrix for Kinect (Group Model)

	Predicted Stuck	Predicted Progress
Actual Stuck	315 (True positives)	144 (False negatives)

Actual Progress	198 (False positives)	243 (True negatives)

None of these results is good when compared to the interaction log algorithm One possible reason is given by analyses of some of the Kinect recordings. We found that some people are calm and do not change body lean when in certain kinds of difficulties, while some people are fidgety and change posture when making progress. This observation and the fact that the individual model gave better results than the group model is consistent with previous results showing that lean differences in posture for the same emotion can be accounted by the differences in subjects' extraversion trait score, and extroverts shift posture more often than introverts (Grafsgaard, Wiggins et al. 2014, Vail, Grafsgaard et al. 2014).

Multimodal Results

As mentioned earlier, in the case of the multimodal approach, we built a group model, even though the lean-based approach gave better results when we used individual models. The confusion matrix in Table VII shows our results of the multimodal approach. In comparison to the log-based approach, the false negative rate went down dramatically from 27% to 7%, allowing us to catch many more difficulties, at the cost of the false positive rate going up slightly from 3% to 5%. Intuitively, the reason for (a) a low false positive rate is that the log-based algorithm was able to filter out the noise from users who were fidgety when they were not having difficulty, and (b) a low false negative rate is that when users were not interacting much with the computer, the body lean was able to compensate for the reduced activity. In terms of accuracy, these results provide the modest gains reported in other studies (D'Mello and Kory 2012) – their importance is based on our application and the rarity of the negative emotion.

Table VII: Confusion Matrix for Multimodal Model

	Predicted Difficulty	Predicted Progress
Actual Difficulty	51 (True positives)	4 (False negatives)
Actual Progress	35 (False positives)	724 (True negatives)

Limitations

Equipment: The Kinect camera had difficulty capturing participants' positions correctly. There are many reasons for this problem such as lighting conditions. Thus, our results are a function of the environment we used. Moreover, the Kinect camera is not standard equipment for programmers and students, which limits their use to helping those who program in special inverted classrooms and

labs. These problems are shared with other multimodal research, which, as D'Mello and Kory (D'Mello and Kory 2012) point out, involves "intrusive, expensive, and noisy sensors ."

Data: Body lean is only one physical feature that can be mined to detect difficulty. Fritz et al. (Fritz, Begel et al. 2014) showed that eye-tracker, an electrodermal activity sensor, and an electroencephalography sensor could be used to predict whether developers would find a code comprehension task to be difficult. The results could be better if we used additional equipment.

Features: It is possible that different features extracted from the tracked data could give better results. For example, a focus-out immediately followed by a focus-in could be classified as navigation to another programming window, allowing other focus events to imply going to an external application to resolve the difficulty.

Choice of machine learning algorithm: We did try another model - a hidden Markov model (HMM) - for the Kinect case, as mentioned in the first author's thesis. The decision tree gave much better results. It may be possible that an alternative model would give better results.

Lab study size and composition: We believe the results from this study do not suffer from over-fitting because (a) the duration of the task was long for a lab study, (b) the results are consistent with our earlier longer-duration and larger lab and field studies of a pure log-based scheme, (c) we used ten-fold cross-validation, and (d) we were not able to get good results by mining facial features or by using a group model for the posture-only scheme. Nonetheless, a larger study would give stronger results. Our "realistic" GUI problem used popular software such as Java and Swing, but it was one created by the experimenters. As in numerous other academic studies, participants in this study were students.

Evaluation: We used the well-accepted Weka's built-in cross-validation mechanism for calculating the confusion matrices and information gain. It would be useful to also do leave-one-out analysis in which we use for testing the data for a specific individual and for training the data for all other participants.

Hypothesized Implications for CSCW

As mentioned earlier, the driving problem for this work is providing unsolicited help to programmers. Here we consider potential relevance of this work in some of the contexts in which such help has been or can be provided. These are hypothesized implications needing, of course, validating future experiments.

Let us first consider software development involving academic or industrial programmers responsible for a common project, a task within the project, or a subtask within the task.

Face-to-face pair programming (Cockburn and Williams 2001): In such programming, one programmer, called the driver inputs a program based on constant consultation with a programmer seated next to the driver, called the navigator, who is expected to be aware of every physical characteristic and action of the driver and thus, does not need any form of computer-supported awareness to determine the difficulties of the driver.

Distributed face-to-face programming (Baheti, Gehringer et al. 2002): This is a variation of the above scenario in which a distributed navigator continuously views the remote desktop of the driver. Arguably, desktop sharing provides all the computer-supported awareness a navigator needs to determine driver difficulties – there is no need for sharing of video or inferences made from it.

Local (Nawrocki, Jasinski et al. 2005) *and Distributed* (Dewan, Agarwal et al. 2009) *Side by Side programming*: As in pair programming, two developers work on the same task and have a persistent audio channel to talk to each other. The difference is that they can work in parallel, using different workstations, on different subtasks of the task given to them. In the local case, they sit side by side so that they can constantly monitor each other's work. In the distributed case, they have a dedicated workstation to show their partner's screen, thereby ensuring that this awareness does reduce the screen real estate available for their own work (Dewan, Agrawal et al. 2010). Pair programming is a special case of side-by-side programming in which only person input at a time. When developers work in parallel, the relevance of our results depends on how much they talk to each other about their problems and how often they monitor their partner's screens and/or persons. The more they talk to each other about their problems, the less relevant our results. When monitoring their partner, our results about feature importance imply that developers should use posture changes of their partners to infer difficulties, and should also look at the workspaces of the partners, especially if the partner is calm. Finally, the cross-validation results imply that difficulty inferences made through our multimodal algorithm are reliable and thus should trigger such monitoring or conversations with the partner to validate them. Dewan, Agrawal et al have run experiments with distributed side-by-programming to determine how often partners do concurrent work, talk to each other, and create conflicts (Dewan, Agarwal et al. 2009, Dewan, Agrawal et al. 2010). While these experiments identify variations based on the pair, they show that the awareness of the other workstation was sufficient to avoid conflicts in all cases. This discussion motivates analogous experiments to determine if workstation awareness is also sufficient to manually infer all difficulties, without the need for automatic inferences.

Radical co-location (Teasley, Covi et al. 2000): Here, developers in a software team are located in a single "war-room" or "bullpen". Our implications here are variations of the ones given above. Team members who can monitor the persons of each other and should use posture changes as difficulty cues. Similarly, they

should use our algorithm for more reliable inferences. Inferences are more important in this case because of the larger group of developers who could be helped (by simply walking over to their seat).

Distributed teams: Here, the team is distributed in different offices, buildings or locations. Again, our work shows the importance of (a) monitoring posture changes (in remote video feeds) and (b) using our inference algorithm. An important difference is that posture monitoring is limited to a small number of remote developers because of the real-estate needed to display remote video feeds. Another important difference is that the cost of investigating a difficulty is high because a virtual collaborative session must be established with the remote collaborator. One way to address this problem is to accompany a difficulty inference with context useful for validating it and determining the nature of the inferred difficulty (Carter and Dewan 2018).

Let us now consider scenarios in which instructors help students with difficulties.

Homework at unscheduled times: This situation, explored in (Carter and Dewan 2015), is identical to the distributed- team case in terms of the implications of our work.

Homework at scheduled times: Here, instructors have allocated time to help remote students, and thus can constantly monitor the work of these students. Codeopticon (Guo 2015) has developed a scheme to support such monitoring. Experience with the system shows that (a) some instructors were overzealous in offering help, and (b) some help offers from the vast majority of instructors were rejected by students. Our work implies that these problems can be reduced by displaying videos and/or difficulty inferences. More important, it shows that instructors do not have to display and constantly monitor workspaces of remote collaborators. Instead, they can perform some background work during this time, and rely on contextualized inferred difficulty awareness to switch to the foreground task of helping students.

Lab-work: This is like the case above, except that there is no need to find real-estate to display the work and/or videos of the students to be helped. The need for reliable manual or automatic difficulty inference is particularly strong here as the cost of missing a difficulty is high in time-bound labs.

Cameras that give joint information are not part of the work environment today. This is why the cues given by viewing the posture changes of collaborators (who can be viewed directly or through video feeds) are important, even though these signals are weaker than the inferences of the multimodal algorithm. Another implication of our work is that controlled public environments such as labs and "war-rooms" should be equipped with such cameras. A related implication is that cameras feeding video to difficulty-inference algorithms or remote helpers should focus on the bodies rather than the faces of programmers.

Conclusions and Future Work

In comparison to other forms of awareness, inferred difficulty awareness offers the promise of opportunistic help that requires less information communication, display, and human processing. Given that, by definition, difficulty is a rare event (otherwise the worker and activity are mismatched), it makes it easier for helpers to find difficulty-needles in the haystack of collaborator-states, thereby making it easier for them to support a larger number of collaborators in difficulty, while doing their own work simultaneously. It does so by using machine-learning to automatically finding these needles. This paper has explored (a) how a state of the art difficulty-inference algorithm should be changed if the haystacks searched include not only the interaction logs but also the videos of the workers and (b) the impact of making the change.

As it is a technical paper, it has focused on the narrow issue of determining the inference to be shared, rather than user-interfaces for sharing this state and triggering opportunistic help giving, which are, arguably, orthogonal issues, addressed by our previous work.

Our work has drawn from previous research in multimodal/multimedia emption detection. It is the only such work that (a) is targeted opportunistic help, (b) addresses programming difficulty, an emotion related to but not the same as confusion, frustration, and code comprehension difficulty, (c) mines a unique and large set of workspace commands, not considered in any previous work on multimodal inference, (d) uses K-mean clustering to classify leans and combines leaning back and forward into one feature in a group model, (e) does an evaluation based on true/false positives/negatives (rather than accuracy) to account for the fact a negative emotion occurs rarely if the task is matched to the subject, (f) shows that log-mining is more effective than posture-mining, though both have high false negative rates, (g) identifies the additive value of mining commands (postures) in addition to postures (commands) and contradicts previous work, based on different metrics (true/false positives/negatives rather than accuracy), that this value is modest (D'Mello and Kory 2012), (h) gathers ground truth by asking programmers to correct predictions of the current implementation of the algorithm in a system they use for everyday work, (i) shows that the additive value of mining postures depends on whether the subject is calm or fidgety, (j) identifies the problems of mining facial features when programmers are actively programming, (k) shows that the individual model worked better in the posture-only case, and (l) does command-based rather than time-based segmentation in the multimodal/multimedia case. Its results strengthen previous findings that show that (a) multimodal effect detection is superior to unimodal detection, and (b) the predictive power of postures depends on the personality of the subject.

As D'Mello and Kory (2012) point out, given the problems of detecting emotions, no one study, no matter how large, can provide the definitive word on the additive value of mining multiple interaction modes. This is why our confirmation of similar earlier findings are perhaps as important as our new ones. This implies that further work is needed to confirm our new results.

Direct observations of the workspace and physical characteristics of collaborators (by observing them directly or through distributed communication channels) give more information than difficulty awareness. For instance, they indicate if the collaborators are frustrated or in difficulty. If such observations are to be replaced (rather than augmented) with inferred awareness, it is important to investigate multimodal inference of other emotions relevant to software development (Dewan 2015).

Moreover, it will be useful to explore the use of integrated log and video-based difficulty/emotion detection in particular and emotion-detection in general to trigger opportunistic collaboration on other kinds of activities such as writing (Long 2016).

This work provides a basis for pursuing some of these promising research directions.

Acknowledgment This research was supported in part by the USA National Science Foundation IIS 1250702 award, an Alumni Fellowship from the Computer Science Department of the University of North Carolina at Chapel Hill, the Brazilian Conselho Nacional de Desenvolvimento Científico e Tecnológico Split Research Scholar Fellowship CNPq process number 203037/2013-8, and an Intel Equipment Award.

References

Abel, M. (1990). Experiences in Exploratory Distributed Organization. Intellectual Teamwork: Social and Technological Foundations of Cooperative Work, NJ, Lawrence Erlbaum.

Baheti, P., E. F. Gehringer and P. D. Stotts (2002). Exploring the Efficacy of Distributed Pair Programming Proceedings of the Second XP Universe and First Agile Universe Conference on Extreme Programming and Agile Methods - XP/Agile Universe 2002 Springer-Verlag: 208-220

Baker, R. (2007). Modeling and Understanding Students' Off-Task Behavior in Intelligent Tutoring Systems Proc. CHI.

Begel, A. and B. Simon (2008). Novice software developers, all over again. International Computing Education Research Workshop.

Calvo, R. A. and S. D'Mello (2010). "Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. ." IEEE Transactions on Affective Computing 1(1): 18-37.

Carter, J. and P. Dewan (2010). Design, Implementation, and Evaluation of an Approach for Determining When Programmers are Having Difficulty. Proc. Group 2010, ACM.

Carter, J. and P. Dewan (2015). Mining Programming Activity to Promote Help. Proc. ECSCW, Oslo, Springer.

Carter, J. and P. Dewan (2018). Contextualizing Inferred Programming Difficulties. Proceedings of SEmotion@ICSE, Gothenburg, IEEE.

Cetintas, S., L. Si, Y. P. Xin, C. Hord and D. Zhang (2009). Learning to Identify Students' Off-Task Behavior in Intelligent Tutoring Systems. Proceedings of the 2009 conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling, IOS Press: 701-703.

Cockburn, A. and L. Williams (2001). The Costs and Benefits of Pair Programming, Addison Wesley.

D'Mello, S. and A. Graesser (2009). "Automatic Detection of Learner's Affect From Gross Body Language." Appl. Artif. Intell. 23(2): 123-150.

D'Mello, S. and J. Kory (2012). Consistent but modest: a meta-analysis on unimodal and multimodal affect detection accuracies from 30 studies. Proceedings of the 14th ACM international conference on Multimodal interaction (ICMI), ACM, New York, NY, USA.

Dewan, P. (2015). Towards Emotion-Based Collaborative Software Engineering. Proc. CHASE@ICSE, Florence, IEEE.

Dewan, P. (2016). Inferred Awareness to Support Mixed-Activity Collaboration. Proc. IEEE CIC. Pittsburgh.

Dewan, P., P. Agarwal, G. Shroff and R. Hegde (2009). Distributed Side-by-Side Programming. ICSE Cooperative and Human Aspects of Software Engineering Workshop, IEEE.

Dewan, P., P. Agarwal, G. Shroff and R. Hegde (2009). Experiments in Distributed Side-by-Side Programming. ICST CollabCom, IEEE.

Dewan, P., P. Agrawal, G. Shroff and R. Hegde (2010). Mixed-Focus Collaboration without Compromising Individual or Group Work. Proc. EICS.

Fish, R. S., R. E. Kraut, R. W. Root, Ronald E. Rice, "Proceedings of the Conference and M. on Computer Human Interaction (CHI) '92, CA, May 1992, pp. 37-48. (1992). Evaluating Video as a Technology for Informal Communication. Proc. CHI.

Fogarty, J., S. E. Hudson, C. G. Atkeson, D. Avrahami, J. Forlizzi, S. Kiesler, J. C. Lee and J. Yang (2005). "Predicting human interruptibility with sensors." ACM Trans. Comput.-Hum. Interact. 12(1): 119-146.

Fritz, T., A. Begel, S. Mueller, S. Yigit-Elliott and M. Zueger (2014). Using Psycho-Physiological Measures to Assess Task Difficulty in Software Development. Proceedings of the International Conference on Software Engineering.

Grafsgaard, J. F., J. B. Wiggins, A. K. Vail, K. E. Boyer, E. N. Wiebe and J. C. Lester (2014). The Additive Value of Multimodal Features for Predicting Engagement, Frustration, and Learning during Tutoring. ICMI, ACM.

Graziotin, D., X. Wang and P. Abrahamsson (2014). "Software Developers, Moods, Emotions, and Performance." IEEE Software 31(4): 24-27.

Guo, P. J. (2015). Codeopticon: Real-Time, One-To-Many Human Tutoring for Computer Programming. ACM Symposium on User Interface Software and Technology (UIST).

Gutwin, C. and S. Greenberg (2002). "A Descriptive Framework of Workspace Awareness for Real-Time Groupware." Comput. Supported Coop. Work 11(3): 411-446.

Harrison, S. and S. Minneman (1990). The Media Space: A Research Project into the Use of Video as a Design Medium. Proceedings of the Conference on Participatory Design.

Hegde, R. and P. Dewan (2008). Connecting Programming Environments to Support Ad-Hoc Collaboration. Proc 23rd IEEE/ACM Conference on Automated Software Engineering, L'Aquila Italy, IEEE/ACM.

Herbsleb, J. and R. E. Grinter (1999). Splitting the Organization and Integrating the Code: Conway's Law Revisited. Proceedings of International Conference on Software Engineering.

Herbsleb, J. D., A. Mockus, T. A. Finholt and R. E. Grinter (2000). Distance, dependencies, and delay in a global collaboration. Proc. CSCW.

Hollan, J. and S. Stornetta (1992). Beyond Being There. ACM CHI Proceedings.

Hudson, S. E. and I. E. Smith (1996). Techniques for Addressing Fundamental Privacy and Disruption Tradeoffs in Awareness Support System. Proc. ACM CSCW.

Isaacs, E. and J. C. Tang (1993). What Video Can and Can't Do for Collaboration. Proc. Multimedia.

Junuzovic, S., P. Dewan and Y. Rui (2007). Read, write, and navigation awareness in realistic multi-view collaborations. Proc. CollaborateCom, IEEE.

LaToza, T. D., G. Venolia and R. Deline (2006). Maintaining mental models: a study of developer work habits. Proc. ICSE, IEEE.

Long, D. (2016). Mixed-focus Difficulty-Triggered Collaborative Writing: Interoperable Architecture, Implementation, and Evaluation, University of North Carolina at Chapel Hill.

Machardy, Z., K. Syharath and Prasun Dewan, CollaborateCom (2012). Engagement Analysis through Computer Vision. Proc. CollaborateCom, IEEE Press.

Mantei, M., R. M. Backer, A. J. Sellen, W. A. S. Buxton, T. Milligan and B. Wellman (1991). Experiences in the Use of a Media Space. Proceedings of CHI'91.

Milewski, A. E. and T. M. Smith (2000). Providing Presence Cues to Telephone Users. Proc. CSCW.

Nawrocki, J. R., M. Jasinski, u. Olek and B. Lange (2005). Pair Programming vs. Side-by-Side Programming. Software Process Improvement, Springer Berlin / Heidelberg. 3792/2005: 28-38.

Tang, J. and S. Minneman (1991). VideoWhiteboard: Video Shadows to Support Remote Collaboration. Proc. ACM CHI.

Tang, J. C. and E. Isaacs (1992). "Why do users like video? Studies of multimedia-supported collaboration." Computer Supported Cooperative Work (CSCW) 1(3).

Tang, J. C. and S. L. Minneman (1990). VideoDraw: A Video Interface for Collaborative Drawing. Proc. CHI.

Teasley, S., L. Covi, M. S. Krishnan and J. S. Olson (2000). How does radical collocation help a team succeed? Proc. CSCW.

Tee, K., S. Greenberg and C. Gutwin (2006). Providing Artifact Awareness to a Distributed Group through Screen Sharing. Proc. ACM CSCW (Computer Supported Cooperative Work).

Vail, A. K., J. F. Grafsgaard, J. B. Wiggins, J. C. Lester and K. E. Boyer (2014). Predicting Learning and Engagement in Tutorial Dialogue: A Personality-Based Model. ICMI, ACM.

Witten, I. H. and E. Frank (1999). Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, Morgan Kaufmann.

Lipovaya, V.; Lima, Y.; Grillo, P.; Barbosa, C. E.; Souza, J. M.; Duarte, F. (2018): Coordination, Communication, and Competition in eSports: A Comparative Analysis of Teams in Two Action Games. In: Proceedings of the 16th European Conference on Computer-Supported Cooperative Work - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2018_11

Coordination, Communication, and Competition in eSports: A Comparative Analysis of Teams in Two Action Games

Viktoriya Lipovaya, Yuri Lima, Pedro Grillo, Carlos Eduardo Barbosa, Jano Moreira de Souza, Francisco José de Castro Moura Duarte
COPPE – Graduate School and Research in Engineering – Universidade Federal do Rio de Janeiro (UFRJ) – Rio de Janeiro, Brazil
vika@pep.ufrj.br, yuriodelima@cos.ufrj.br, grillo.pedro@pep.ufrj.br, eduardo@cos.ufrj.br, jano@cos.ufrj.br, duarte@pep.ufrj.br

Abstract: eSports are increasing in popularity and in importance worldwide and, given that they essentially involve the cooperation of teams competing among themselves, they are an interesting study object for the CSCW field. In this study, we contribute to the CSCW literature regarding eSports by performing a comparative analysis of two different action games, focusing on how cooperation, communication, and competition take place in each one of them. To do so, we perform a semi-qualitative study involving interviews with professional and amateur players. Then, we analyzed the results of the fieldwork, which consisted of a 31-question questionnaire with 65 valid respondents. Moreover, we discuss and highlight the relationship between our results and other CSCW-related works focusing on our research questions. Among our findings, we can highlight the specialization of work in different eSport teams, the importance of non-verbal communication during matches, and the interplay between competition and collaboration in the same team.

Introduction

Over the last decade, online games have become the industry standard. One segment of this industry — competitive games — is becoming increasingly popular. A competitive game is one in which players can compete against each other, either by themselves or in groups. The popularization of competitive games has enabled the creation of specialized events, usually broadcast via the internet, in which teams compete in championships. The prizes offered, both financial and non-financial, have led to the creation of professional teams, specialized in a given game. The concept of electronic sports (eSports) emerged from competitions involving professional players.

There is substantial financial potential in eSports given that the revenue in 2016 was US$492.7 million, with a forecast of US$1.5 billion for 2020 (Newzoo 2018), which indicates steady growth. Initiatives to take eSport content to traditional media, such as TV, are being developed (Beck 2017a, 2017b, Bullock 2017), and efforts are being made to make eSports comparable with traditional sports, in the hope of inclusion in the Olympics (Graham 2017, Good 2017). Several characteristics of eSports are challenging themes in the study of collaboration, given that players can be geographically dispersed and must execute a task — in the group — that demands constant decision making and the real-time sharing of situational knowledge.

Computer-Supported Cooperative Work (CSCW) is a collaboration-themed research field that seeks to understand how people work in a group to attain a common goal through the use of computational tools (Wilson 1991). We adopt the paradigm that an eSport game would be an interesting research subject for CSCW — as proposed by Freeman & Wohn (2017b).

In this study, we focus on analyzing the collaborative aspect that arises in eSport teams of high-performance players, highlighting the differences between amateur and professional players. In particular, we try and fill a gap in the CSCW literature regarding eSports as indicated by (Freeman and Wohn 2017b)): a comparative analysis of the collaboration that occurs with players of different games. Thus, in this study we answer three research questions:

- RQ1 — How is work coordinated in eSport teams?
- RQ2 — How do players communicate during matches?
- RQ3 — Are there conflicting interests during the matches that make players in the same team compete against each other?

In order to answer these questions, we did a qualitative study using structured and semi-structured interviews with 74 eSport players.

This study is divided as follows: in section 2, we analyze and discuss the current state of eSports; in section 3, we discuss how collaboration (specifically CSCW) is related to eSports; in section 4, we detail the methodology used in this work; in section 5, we present the results obtained; in section 6, we discuss the

results in light of the CSCW literature; in section 7, we present our conclusions; and in section 8, we discuss future works.

The Current State of eSports

eSports have been growing in number and in terms of the academic debate, which can be seen in the plurality of academic definitions regarding the concept (Freeman and Wohn 2017a). In this study, we adopted the definitions most relevant to the CSCW area. Wagner (2006) and Hamari & Sjöblom (2017) define eSports as computer-mediated sports and an area of sports activities in which people develop and train physical and mental abilities using Information and Communication Technologies. According to Freeman & Wohn (2017a), within the research areas of CSCW and HCI that concern the study of eSports, most authors (McClelland et al. 2011, Hamilton et al. 2012a, 2012b, Kaytoue et al. 2012, Kow and Young 2013, Leavitt et al. 2016) describe eSports as competitive computer/online games. The competitions can involve several levels and scopes, from a local match using Local Area Networks (LAN) to national and international championships.

As reported by (Newzoo 2018), the eSport industry is a significant market that won't stop growing. Having produced US$492.7 million in revenue in 2016, and with an annual growth of 34%, this market is expected to reach US$1.5 billion in 2020. Another financial element that shows how the eSport industry is already huge, is the value of the prizes of the highest paying championships. Currently, more than 40 competitions have a prize pool over US$1 million, and The International 2017 (T17) — a Dota 2 championship — is the biggest competition in terms of prize pool, with US$24 million in prizes.[1]

Several authors have discussed the relationship between traditional and electronic sports (Skubida 2016, Freeman and Wohn 2017a, Jenny et al. 2017, Funk et al. 2017, Hallmann and Giel 2017, Heere 2017). Like any professional athlete, professional eSport players have a training routine, an agenda full of events, and they must maintain a healthy relationship with supporters and sponsors. These similarities in traditional and electronic sports are currently very relevant, given the institutionalization possibilities worldwide. The International Olympic Committee announced that eSports would be medal events at the 2022 Asian Games in China, and it is also possible that they will be included in the 2024 Olympic Games (Graham 2017). Also in this trend of professionalization, in the US, professional players can request the P-1 visa for athletes (Academy 2017). In Brazil, the Senate is analyzing a bill for the regulation of eSports (Senado Federal 2018).

[1] https://www.esportsearnings.com/tournaments

eSports and CSCW

eSports involve collaborative virtual environments and a complex social organization between players (Brown and Bell 2004), which makes them relevant to the CSCW community. There is some research on eSports in the CSCW and HCI literature, which includes: measuring the attentional and cognitive abilities of eSports players, in order to differentiate them from non-videogame players; estimation of the necessary effort to become an elite player; using of game concepts in the design of collaborative systems (gamification); and understanding the social dynamics in gameplay (Kozachuk et al. 2016). According to Dafai (2016), eSports can be similar in design despite their genres. Dafai identified five design characteristics that League of Legends (LoL) and Counter-Strike: Global Offensive (CS:GO) have in common, and may be seen as *essential* for a successful eSport. The design characteristics that these two eSports explicitly share are *Match Based Structure*, *Player Evaluation System*, *Explicit UI*, *Player Performance Feedback* and *Game Client* (Dafai 2016).

Taylor (2012) deeply analyzes eSports – comparing it to regular sports – exploring how gamers become professionals, how eSports are structured and its culture, how the global and local contexts affect eSports, and discusses if eSports are a serious leisure or a true form of professional play.

Several eSport games are designed to support collaboration in competitions between teams of players, which makes them suitable with both the 3C collaboration model (Fuks et al. 2008a) and the concept of awareness (Teruel et al. 2016). In this section, we will analyze how cooperation, coordination, communication, and awareness apply to eSports, using the mapping of the 3C model for adaptive workflows, which are similar to the one in eSports (Fuks et al. 2008b).

Communication is used in the decision concerning the distribution of activities among team members, and to synchronize and renegotiate this division depending on the situational context (Fuks et al. 2008b). According to Leavitt et al. (2016), due to the frantic rhythm and the ad hoc nature of communications in eSports, the primary means of communication used are voice and text. The use of the "ping" — a non-verbal communication that marks, in the virtual environment, a situation to be acknowledged by the members of a team — is another option for improving situational awareness that has a reduced impact on the focus of the team members when compared to verbal communication (Leavitt et al. 2016). Not every eSport game has the ping as a communication method, but game dynamics, in general, tend to allow for such non-verbal communication, as shown by (Toups et al. 2014).

Coordination is responsible for the breakdown of the main goal into different activities (partial goals) and the distribution of these activities among the members of the workgroup (Fuks et al. 2008b). The execution of these activities in an integrated manner is responsible for the performance obtained and goals

achieved (Freeman and Wohn 2017b). In accordance with Freeman & Wohn (2017b), the mental models shared regarding the tasks to be performed as well as the interaction between the team members allow for anticipation and prediction of the behavior of the team's companion — an emergent phenomenon called *team cognition*.

Cooperation can be defined as the execution of distributed activities (Fuks et al. 2008b). Lameiras et al. (2014) argue that there are two types of cooperation for athletes of traditional sports: conditioned cooperation and unconditioned cooperation. Conditioned cooperation is related to the perception that the athlete can achieve his personal goals through cooperation with the team. In the unconditioned cooperation, cooperation happens regardless of the personal goals of each athlete. Lameiras et al. (2014) also indicate that situational factors can induce cooperation.

The awareness of the team, in turn, is generated and mediated by communication, coordination, and cooperation (Fuks et al. 2008b). Endsley (1995) argues that situational awareness (individual) involves the perception of the elements in the present moment, the understanding of their meaning, and the projection of the situation in a near future. Situational awareness is fundamental to the decision-making process of the actions that will be immediately taken in an eSport match to achieve a given goal.

Freeman & Wohn (2017b) emphasize that the study of eSports offers an opportunity for discussions in the area of CSCW, given that eSports feature hybrid collective work. eSport teams are a mix of two types of teams — high-performance teams and decision-making and knowledge-intensive teams. The teamwork involved in them occurs in highly competitive, stressful, and intense virtual environments that demand fast decision making and action taking associated with physical (both virtual and non-virtual) activities. Thus, these teams are oriented toward action, particularly in the action games that are the focus of this study.

Given this theoretical context, this present study was guided by the following characteristics of the definition of an eSport: (i) computer-supported cooperative work; (ii) involves the collective aspect of work; (iii) a physical and intellectual activity undertaken during training or a championship; (iv) has as a goal the defeating of opponents; and (v) represents a hybrid type of virtual team.

Methodology and Approach

We present below the summary of our research design, which follows (Maxwell 2009) model, with five components:

(1) *Goals*: In a theoretical aspect, we want to contribute deepening recent discussion regarding eSports and coordination and communication, by focusing on action games. In practical terms, we want to contribute

exploring subjects that are rising in this emerging field, in order to help future research and policies that can be made.

(2) *Conceptual framework*: As described in the previous chapter, we covered eSports and CSCW research, but we focused on using a recent study from Freeman and Wohn (2017b) to drive part of research concepts. Besides CSCW, another conceptual framework of our study came from our background in Ergonomics, which motivated our feedback loops to validate some research issues with professional players.

(3) *Research questions*: We focused on three research questions: RQ1 – How is work coordinated in eSports teams? RQ2 – How do players communicate during matches? RQ3 – Are there conflicting interests during matches that make players in the same team compete against each other?

(4) *Methods*: We used a 31 item online-based questionnaire (see Appendix), based on previous research (Freeman and Wohn 2017b), with a total of 65 final responses. Our approach in the process of elaboration, validation and data collection and analysis is described in Figure 1. The responses profile is shown in **Table 1**.

Table 1. Demographic profile of the respondents (N=65)

Sex	Female: 5 (8%) Male: 60 (92%)
Age	Mean: 19.8 years Oldest: 36 years Youngest: 12 years
Country	Brazil
Category	Professionals: 14 (22%) Amateurs: 51 (78%)
eSport games played	Counter-Strike: Global Offensive (CS:GO) — 25 (38.46%); Rainbow Six: Siege (R6) — 23 (35.38%); and Others — 17 (26.16%), which includes League of Legends (LoL) — 5, Overwatch — 5, Fifa — 4, Clash Royale — 2, and Dota 2 — 1

(5) *Validity*: We used some feedback loops in order to increase questionnaire and responses coherence and validity, by confronting them with experienced players and professionals in eSports. It is important to remember that it was conducted only with Brazilian players and with a short period for data collection for this exploratory paper. For further investigations, it might be necessary a more complex and cross-sectional study.

The approach of the research followed three main phases, which are summarized in Figure **1**, and briefly described in this section:

I. *Exploration*: Some of the authors of this work have done interviews with eSports professionals in the past during the course of conducting other research (Lipovaya et al. 2017, Ikenami et al. 2018), this ongoing research connected with CSCW literature and a research group was formed with experiences regarding eSports and CSCW. Some interviews were made with four eSports experts in order to validate literature questions and explore problems, this helped us elaborating the data collection procedure, including the questionnaire.

II. *Data collection*: The final questionnaire was validated and evaluated with other eSports professionals. In order to get more responses, we made an engagement strategy with eSports influencers to share the questionnaire using social media.

III. *Data analysis*: Data analysis and compilation of the results involved a three-level structure – the highest level is the research question, followed by the type of the game, and, finally, player class (amateur or professional). We used a qualitative analysis methodology, first individually finding patterns in responses and then collectively constructing results, using the data analysis hierarchical structure.

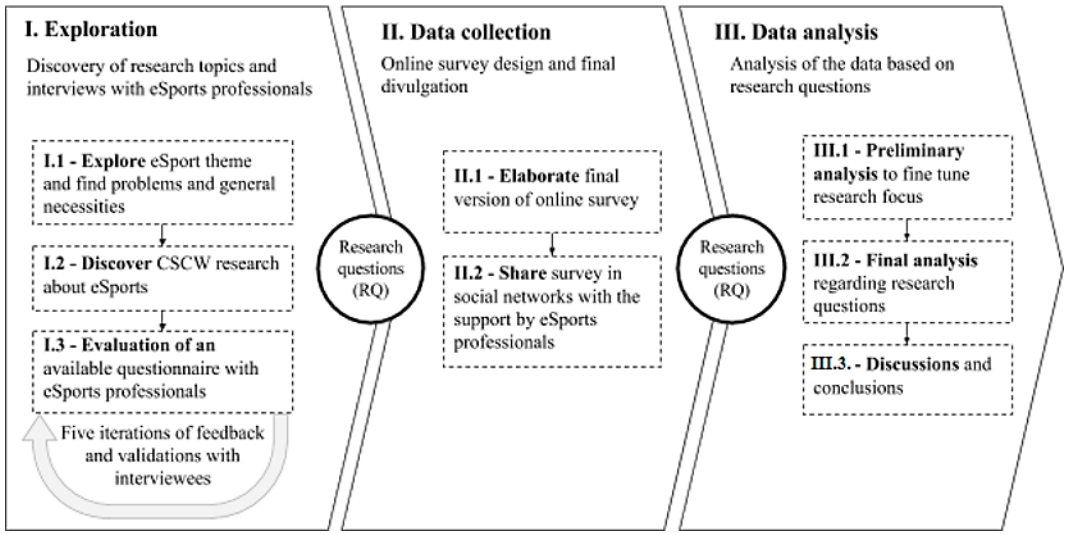

Figure 1: Summary of the phases and steps of the methodology.

Analysis of Results

Coordination

Given that understanding the division of work is essential for understanding how coordination is achieved (Mintzberg 1989), we sought to explore this question that remains underexplored in the literature on eSports. Thus, our field research was aimed at understanding how work is divided both vertically and horizontally; and understanding what is expected of each role in the team.

During the exploratory phase of our study, we found that, besides players, there are three important roles in eSport teams: captain, coach, and manager. In the questionnaires, we set out to understand if each of these roles exists in eSport teams, and if they are roles assumed by a single person (e.g., a player that is also a manager), and we also sought to understand what is expected of each of these team roles during the matches.

Captain

Overall, 62% of the interviewees said that the role of captain exists and that the captain plays with the team. Only 5% said that the captain exists and does not play with the team, while 29% said that there is no specific captain role, and of this percentage, 13% said that the role of the captain is divided between the team, as shown by Figure **2**. For the CS:GO players, nobody said that the role of captain is taken by a non-player, and 67% affirmed that the captain is a player. Among the R6 players, this percentage increases to 78%. From these answers, we concluded that the role of captain exists and it is either a responsibility given to one of the team players or distributed among them — in other words, the non-player captain is uncommon.

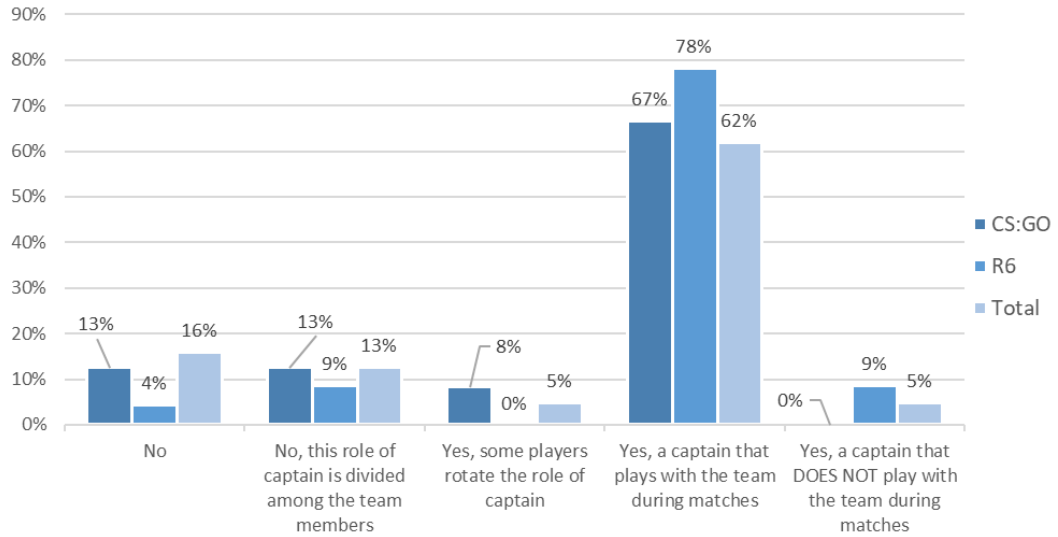

Figure 2. Answers to question 13 — "Is there a captain in your team (someone that leads the team during a match)?" — categorized by game.

Upon comparing professionals with amateurs, we realized that in the first group, nobody said that the role of the captain is divided among the players; whereas 16% of the amateurs said that this happens. Among the professionals, 17% said that there is a captain who does not play with the team, especially those who play R6, which may indicate a pattern in this group, considering that two out of the three respondents said that a captain is a person who is not part of the team. Among the amateurs, just one person gave this reply. Meanwhile, all the CS:GO professionals said that the captain is one of the players on the team. Therefore, we can conclude that there is a difference between the professional and amateur eSport teams with respect to the role of the captain, given that the professionals do not divide this role among the players in the team. Another conclusion is that professional R6 teams differ from the amateurs and CS:GO professionals because the role of captain tends to be given to a person who is not a player.

Concerning what is expected from the person who assumes the role of captain during a match, the analysis of the answers allows us to infer that the captains of the CS:GO and R6 teams must remain calm, define the tactics, and morally support and motivate the team.

Coach

When asked about the existence of a coach, the replies indicated no relevant differences when comparing the games. The only interesting facts are that in CS:GO teams there is more likely to be a coach and it is less likely that the coach plays with the team. On the other hand, there were some relevant differences when comparing professional and amateur players. As can be seen in Figure **3**, 91% of professionals said that the role of coach exists, and 55% claimed that the coach is not one of the team's players. Among the amateurs, 60% said that the role of a coach either does not exist or is divided between the players — no professional gave the latter answer. The responses tell us that, in general, the professional teams have a coach — a trend that is not surprising, given the competitive nature of the industry and the financial interests vested in the victory of these teams. However, the role of the coach is uncommon among amateurs, who rarely have a coach who is not part of the team.

During a match, it is expected that a coach of a CS:GO team will provide tactical support with his knowledge, review the team's mistakes, and help the team members with issues not related to the game itself. In R6 teams, the coach must study the adversary and the maps to formulate strategies for the team. In both games, the coach also has to perform some activities — such as motivating and leading the team — that are similar to the role of a captain.

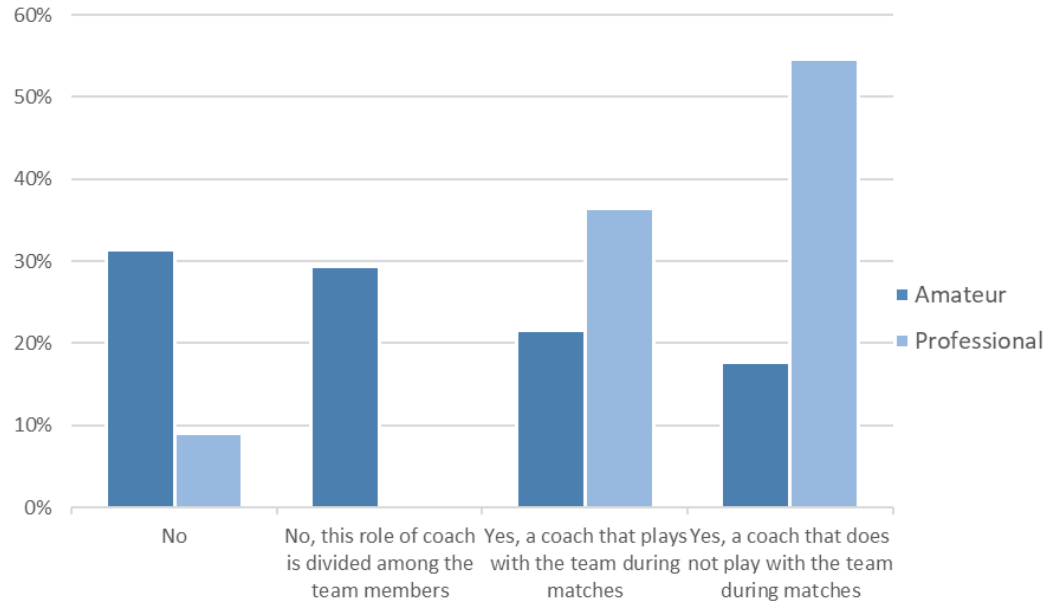

Figure 3. Answers to question 15 — "Does your team have a coach?" — categorized by player class

Manager

The results of the field research showed that the existence of a manager is much more common among professionals (69%) than among amateurs (37%). Comparing the games, we could see that 71% of the respondents who play CS:GO said that this role does not exist in their teams; while 56% of the R6 players said that someone has the role of manager, but the manager is usually (39% of the replies) not one of the players.

eSport players expect that during the matches, the manager will, if needed, give some support in matters unrelated to the game itself. The manager is more active outside the matches, mainly taking care of issues related to sponsors, registration in championships, and team marketing.

Player roles and role rotation

As seen in Figure **4**, R6 and CS:GO teams usually have a different number of roles. The CS:GO teams have more roles than the R6 teams — five roles was the most common (52%) for CS:GO teams; while for the R6 teams, three roles (also 52%) was most common.

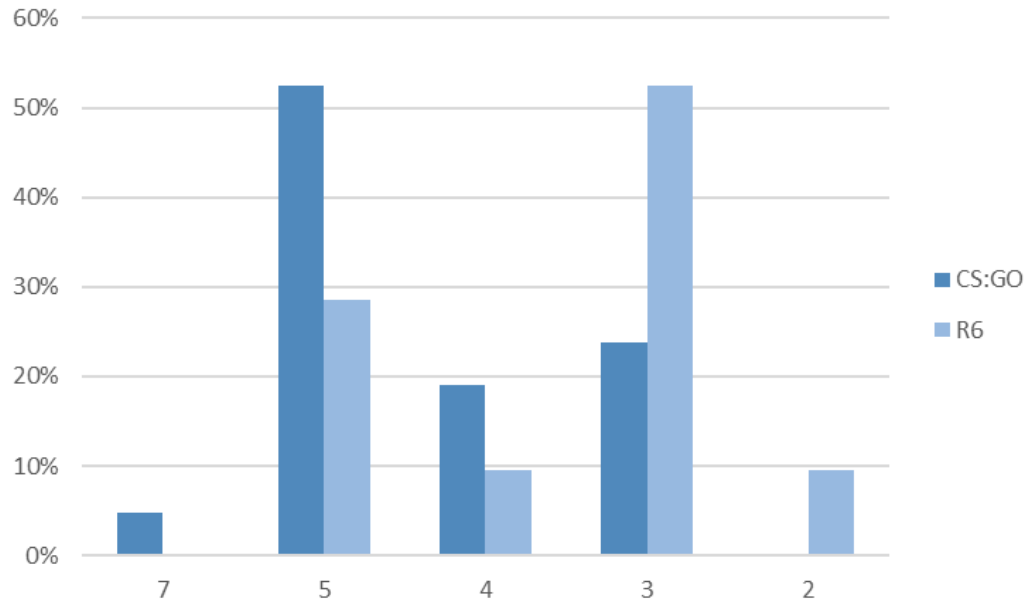

Figure 4. Number of roles in a team, categorized by game.

Among the CS:GO players, the most frequent role was Entry Fragger, which was cited by 100% of the respondents, followed by Support (95%), and AWPER (or Sniper) with 77%. Meanwhile, the roles that R6 players cited most were Support (100%), Entry Fragger (67%), and Intermediary (62%).

Role rotation is a common practice among the eSport players who answered our questionnaire — 82% said that they have changed roles at some time in the past. For the R6 players, in 50% of the cases, these changes happen between championships, with change during a championship (13%) or a match (23%) being less frequent. On the other hand, in the CS:GO group of respondents, role rotation is more common during a match (50%) than between championships (23%). Considering the professionals of these games, we can see that the trends are very different — for the R6 players, only one interviewee said that role rotation happens between matches, while the other two said that it does not happen at all; however, of the five CS:GO professionals, three said that role rotation happens during matches in progress.

Communication

In this section, we will analyze how eSport teams communicate. Among the means of communication, the use of voice chat is universal, followed by the use of in-game visuals, text, real-world visuals, and videos, as shown in Figure **5**. When comparing professionals and amateurs, we can see that professionals do not use real-world visual communication in matches, and this is also uncommon among amateurs. Upon analyzing Figure **5**, we can see that R6 players tend to use the most number of different types of communication within their teams.

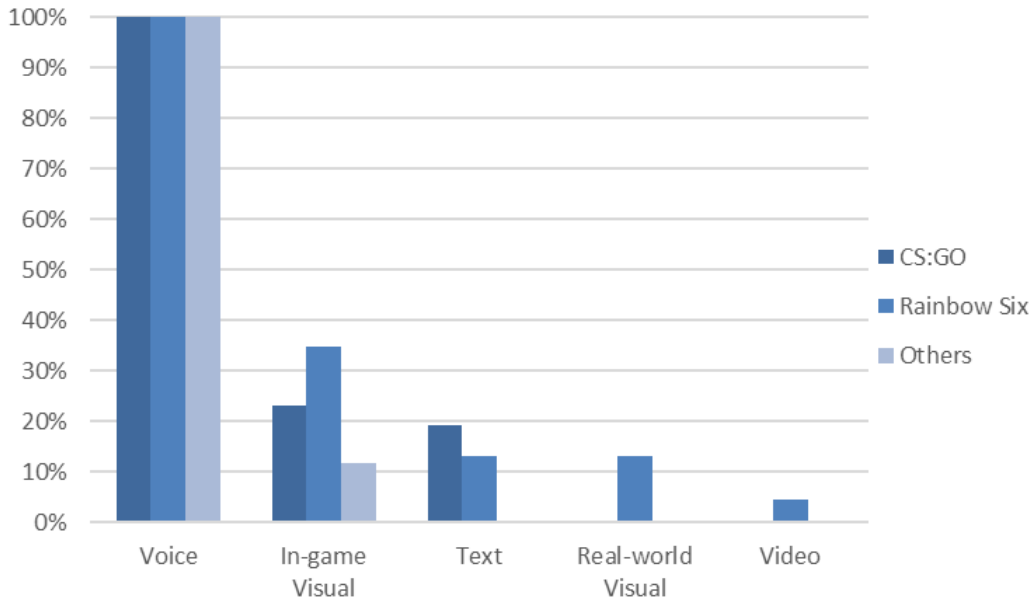

Figure 5. Type of communication used by interviewees, categorized by game.

We can see something relevant when analyzing the tools used to support voice chats: professionals use only one tool. On the other hand, 34% of amateurs use more than one chat tool. Another observation we can make is that — as shown in Figure **6** — the choice of communication tool appears to be segmented according to the game. CS:GO has a higher number of users of Teamspeak than the mean values for the other games. Nevertheless, Discord is especially popular among players grouped in the "Others" category, and Party is a tool commonly used by R6 players.

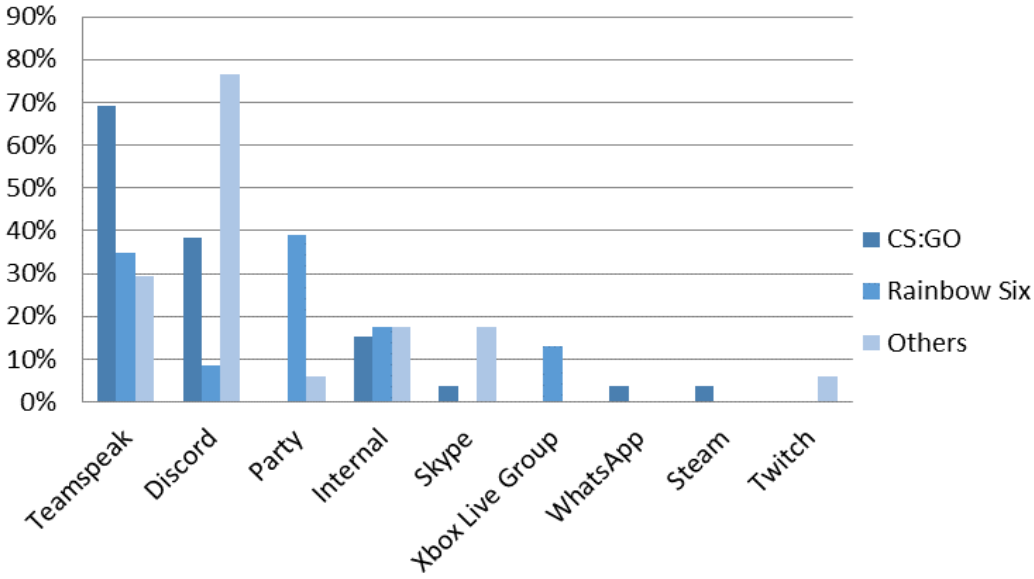

Figure 6. Tools used by interviewees to support voice communication, categorized by game.

Before and after-match activities

Here we analyze the activities that teams do before and after matches. For the activities done before matches, we have the following classification:

- Warm-up training: training with a focus on preparing the players for the upcoming match;
- Reviewing strategies/tactics: discussing the techniques that will be used and the responsibilities given to the players;
- Adversary focus: discussing how the opposing team plays and determining strategies to overcome its playing style;
- Generic chat: "small talk" among team members;
- Concentration: the team just seeks to focus its attention and energy on the forthcoming match;
- Watching matches: the team watches some matches of the opponent together.

Analyzing the data collected, which is shown in Figure **7**, we can highlight that professional players tend to review strategies/tactics and focus on studying the adversary more than amateurs do. When comparing the games, it can be seen that warm-up training is more common among R6 players (more than 50%) than CS:GO (less than 30%).

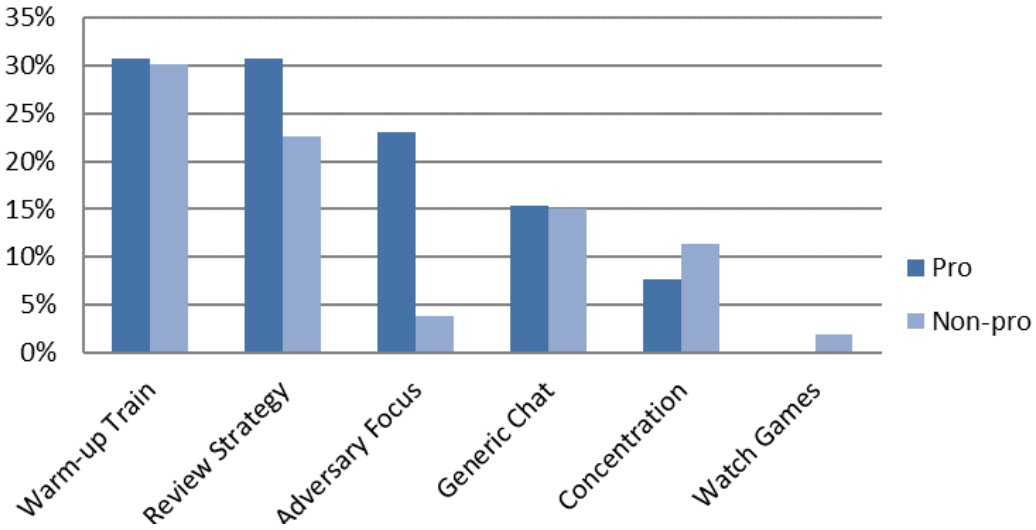

Figure 7. Activities performed before the game, categorized by player class.

Considering the activities performed after the match, the answers were categorized as follows:

- Post-match review/chat: critical analysis of the match played, with the aim of identifying and correcting errors;
- Relaxation: moment dedicated to relaxation;
- Training: playing other matches (ranked or not) in order to improve skills in a given game
- General chat: "small talk" among team members.

Upon examining the activities performed after the match (Figure 8), we can see that the most common activity among professionals and amateurs alike is the post-game review/chat. One relevant difference between these two groups is that relaxing after a match is more common for amateurs (37%) than professionals (15%). Upon comparing the different games, we can see that 30% of R6 players train after matches, something which is uncommon for the CS:GO players (less than 5%) and "Others" (10%).

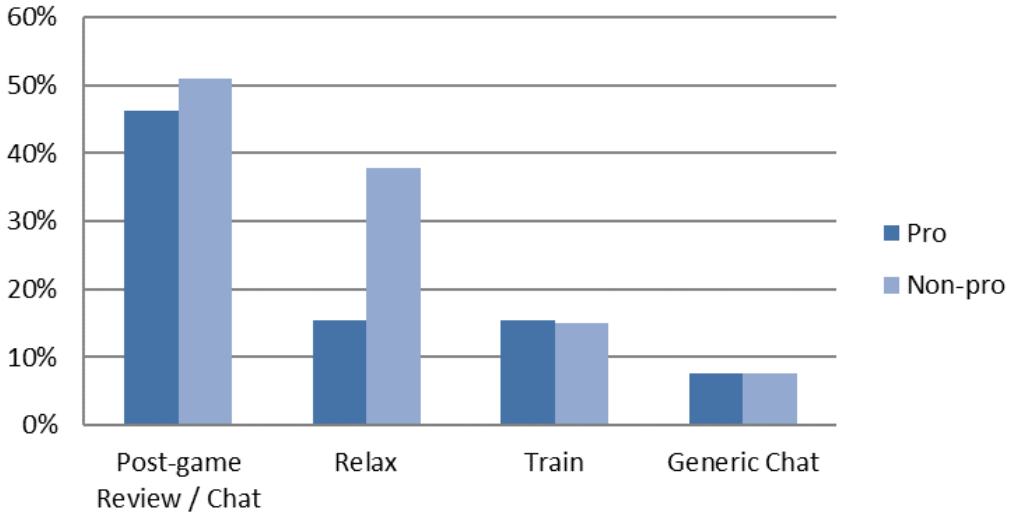

Figure 8. Activities performed after a match, categorized by player class.

Collaboration and competition

This section will present the analysis of the answers given by the players about how collaboration and competition occur in their teams. Firstly, we will present the results of the comparison of the answers regarding collaboration, followed by the analysis of the competition within the teams.

Collaboration

The analysis of the answers to question number 25 — about collaboration during matches — was based on the main keywords that players used to describe the collaboration situations. Most of the respondents (55%) described collaboration as being tasks that are part of their work helping teammates, and they discussed examples in which collaboration occurs in their teams, using game terminology such as "throwing bombs", "killing enemies", "giving help to jump", and "covering someone". One example is described by P20 (male, 15, professional): "[I collaborate] when I have to destroy an enemy gadget with a shock drone or incapacitate a colleague so that he can come back with more HP (Health Points)".

Some players (19%) answered the question by speaking more broadly about collaboration as the importance of teamwork, the creation of strategies, and rehearsed plays, as commented by P34 (male, 29, professional): "I have to help my partners to take others out of position".

When compared with amateurs, professionals emphasized the importance of training that helps in the development of individual and collective strategy — especially attempts to improve tactics in order to improve collaboration, as stated by P8 (male, 26, professional): "I help the captain at certain moments so that we can finish the match with the right strategy". Additionally, when talking about

collaboration, professionals mention the importance of the gathering and transfer of information regarding what happens in the game, in order to always keep their teams up to date.

Competition

When asked if there is competition within their own team, 42% of all the respondents said that it exists — 62% of the professional players, but only 38% of the amateurs. Thus, there is a huge difference between the perception of professionals and amateurs when talking about competition.

In the explanations of the reasons for internal competition, 41% of players said that it is for recognition (e.g., being chosen as the most valuable player — MVP), 17% said it is just for the fun of competing, and 14% stated egotistical reasons — see Figure **9**. One of the main differences observed between amateurs and professionals was that egotistical reasons are more common among professionals (25%) than amateurs (10%). Additionally, reasons for competing — such as discomfort with the team, motivation, and disputes between starters and substitutes — only appeared in the amateurs' answers. The comparison between the different games did not show relevant differences.

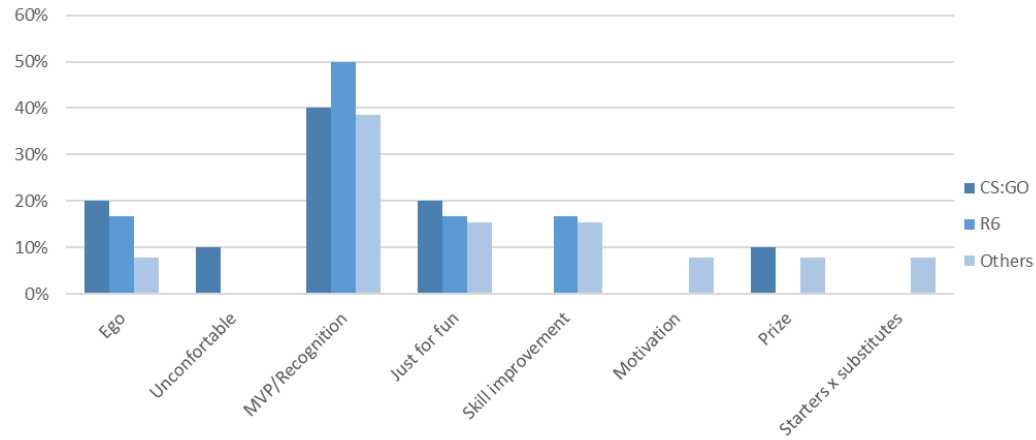

Figure 9. Players' reasons for competing, categorized by game.

Discussion

Having presented the analysis of the results of the field research, we will now conduct a preliminary discussion about the main topics that are relevant for the CSCW field, focusing on the research questions previously introduced. We concentrate our analysis on the two games with the most responses to our questionnaire (CS:GO and R6). Besides, we were able to reach professional players of these two games to give some of their insights about the results of our study, further improving our discussion.

Coordination

The fact that eSport teams — especially those that play action games — can be considered to be a mix of action and knowledge-intensive teams — as argued by (Freeman and Wohn 2017b), lets us compare eSports with other types of collaborative work — such as trading, transit control, and shipping navigation — previously explored by CSCW literature (Cheung et al. 2012).

As we explored action games in this study, it was clear to see how eSport teams can be considered to be action-oriented, as classified by (DeChurch and Mesmer-Magnus 2010). Both of the games are first-person shooters (FPSs), which means that players see the virtual world through the eyes of their avatars, making immersion in the action bigger than third-person shooter (TPS) games. Additionally, both games allow players to customize their avatars and weapons, further improving the usual identification of players with their virtual representations (Livingston et al. 2014). These factors give players such an immersion in the action as well as identification with their virtual selves that we can speak of eSport teams playing action games as virtual action teams.

When discussing the vertical specialization of work; that is, the separation of the execution of work from its administration (Mintzberg 1989), our results showed that the captain and the coach are the roles that the players expect the most from during matches; whereas managers' responsibilities involve helping the team before and after matches with matters such as registration, marketing, and public relations. Therefore, managers are not as important as the captain and coach during the actual competition. The role of captain exists in different ways — a player who performs this role was the most common in our findings, and it is expected that the person who has this role will define the tactics to be used by the team. This could be due to the fast-moving environment of eSports, which provides less opportunity for players to communicate and consider everyone's opinion (Kim et al. 2017).

The other type of work specialization — horizontal specialization or the division of parallel activities (Mintzberg 1989) — was also explored in our fieldwork. In order to discuss the results, it is necessary to understand what roles are available in the action games explored in our research. Despite being action games, CS:GO and R6 have several differences; therefore, we sought a professional player of each game (who had responded to our questionnaire) to help us understand these games, and we also consulted the fan-created wikis for each game.[2,3] CS:GO is an FPS featuring an armed confrontation between teams of terrorists and counter-terrorists. R6 is a tactical FPS that involves armed close combat between two teams of counter-terrorist groups. Due to these differences in

[2] http://counterstrike.wikia.com/wiki/Counter-Strike:_Global_Offensive — Accessed on January 5th 2018

[3] http://rainbowsix.wikia.com/wiki/Tom_Clancy%27s_Rainbow_Six_Siege — Accessed on January 4th 2018

the goal of the game, CS:GO and R6 teams have different roles. Even roles that have the same name in both games (e.g., *Support* and *Entry Fragger*) are essentially different. To give an overall understanding of the variety of roles that may exist in an action game, we will describe the roles in CS:GO that were most cited by interviewees. An *Entry Fragger* is the first combatant that seeks information behind enemy lines and tends to be the one with the most kills. The *Support*, as the name implies, is expected to give some support to *Entry Fraggers* by playing more defensively. An *AWPer* or *Sniper* is a player who has a powerful weapon (that can kill with only one shot) and can help both defensively and offensively.

Most players have changed roles at one time or another in the past, which could be a factor in improving the collective intelligence of a team (Kim et al. 2017) since this gives players a better understanding of everyone's responsibilities during a match. This rotation of roles or tasks is a normal business practice that provides several benefits to employees; for example, enhanced career development (Campion et al. 1994) and increased versatility (Eriksson and Ortega 2006).

Communication

As seen in the analysis of the results, players tend to use different methods to communicate with their teammates. One of these methods is non-verbal communication (cited by 25% of the players), which means that it is a relevant communication method. Thus, as previously shown by (Leavitt et al. 2016) in LoL (a strategy game), the importance of in-game visual communication was confirmed by our study of action games, and we inferred that this type of communication is important for eSports in general, and the way that it is employed by teams should be further evaluated. We believe that the use of non-verbal communication is even more common than reported by the interviewees because it is common knowledge in the CSCW literature that actions can be a substitute for verbal communications in shared visual space (Gergle et al. 2004). However, players have difficulties in externalizing the use of this communication method, because it is deeply rooted in their tacit knowledge and work practices.

As stated by (Cheung et al. 2012), non-verbal communication is even more important in fast-paced collaborative games like FPSs (i.e., the games we analyzed). The quick and ad-hoc decision making of action games demands that players find ways to communicate faster than verbal communication (either voice or text), just like or even more so than for strategy games.

As (Leavitt et al. 2016) showed, non-verbal communication methods improve the situational awareness of teams but can also interrupt a player's flow, disrupt their focus, or overload their attention. The fact that only 15% of professionals cited in-game visuals as a communication method might indicate that these more experienced and serious players understand the negative effects related to the use

of these methods and prefer to use verbal communication in order to avoid this issue during matches. Moreover, professionals tend to know better how they must act within a team, as (Mason and Clauset 2013) argue, which makes them less dependent on communication in general.

Competition

The fact that players who are part of a given eSport team have to collaborate within a game, makes competition natural — gamification (even more so for games that depend on cooperative-competitive features) tends to cause competition even in environments in which it is undesirable (Morschheuser et al. 2017). Also, given that eSports are similar to traditional sports in several aspects, including their competitive nature (Jenny et al. 2017), it is expected — and our field research shows this — that there should be competition among players in the same team, because they want to be the MVP (same term as used in traditional sports). In a still growing market, particularly in Brazil, being the MVP allows players to be noticed by scouts and hired by bigger and better-paying teams.

As question number 35 of our questionnaire allowed respondents to leave their contact information if they so desired, in order to further improve our understanding of the competition within eSport teams, we decided to contact two of the respondents: one CS:GO player and one R6 player, but both professional players. The results of the questionnaire indicated the existence of a competitive climate inside the teams, whether amateur or professional. This can be explained by a desire to be the best and to improve oneself, which is one of the main goals of any athlete. When compared with amateurs, the professionals indicated competition as a need to guarantee their salary and acquire other sources of income (such as sponsorships), ultimately allowing them to dedicate themselves entirely to eSports as a job. As explained by the professional CS:GO player:

> eSports is an intrinsically competitive modality — competition is part of success. The goal of a team is to win collectively. However, the goal of each athlete is to be the best individually, rising up the MVP list.

Thus, in eSports we can see a mix of conditioned cooperation and unconditioned cooperation (as defined by (Lameiras et al. 2014)) as, even if the desire to achieve high individual performance exists, eSport players need to understand the importance of collective work, which is impossible if the players do not help their teammates in the execution of their activities as well as fulfilling their own roles. As explained by the professional R6 player contacted:

> For a team to win, players must be able to collaborate among themselves, balancing the aspects of cooperation and individual competition. It could be a tough day for the ‘star’ of the team… so what? The team cannot let its level drop,

and the game is too dynamic, everything happens quickly, so players must be quick on their feet. Not only speed of reaction, but an understanding of the game is also needed.

Final Remarks

In this exploratory work, we set out to answer three research questions, which are repeated here for convenience:

- RQ1 — How is work coordinated in eSport teams?
- RQ2 — How do players communicate during matches?
- RQ3 — Are there conflicting interests during the matches that make players in the same team compete against each other?

Regarding the first question, our results focused on the division of work in the eSport teams. We were able to understand what the main roles are, and what is expected of each one of them during matches, which lets us see how two different action games can be in terms of coordination. The second question involved the exploration of the main methods and tools used by the players on a given team to communicate during matches. Our research showed the importance of verbal communication followed by in-game mechanics. In relation to the third question, we were able to confirm the existence of competition between players within the same team, which is usually motivated by an interest to be recognized as the best player (MVP), and to improve oneself — something natural among athletes in general.

Our research also indicated some improvements for the design of games played in eSport competitions; for example: giving players more options to use non-verbal communication, while considering the distraction that this method may cause; and looking at how coordination is undertaken in real-world action teams in order to use this knowledge for game mechanics. As games — which were initially seen as essentially entertaining artifacts — are increasingly used for collaborative and competitive work, areas such as HCI and CSCW need to develop more research to support the design of eSport games, and we believe that our research is one of many small steps in this direction.

This study, which is part of ongoing research, certainly has some limitations. The time period for the data collection was short, yielding only about 80 fully answered questionnaires, thus reducing the robustness of our analysis. We interviewed some of the respondents personally to improve our discussion of the results, but due to time constraints, we were unable to fully process the results of this step, leaving it partly out of the text for this study.

Future work

This study is being expanded by personally interviewing people who perform the other roles mentioned in the course of the work: players, coaches, captains, and managers of professional eSport teams. Another method that we plan to use is the observation of teams' coordination, communication, and competition during the championships. We want to discover how the people with different roles within these teams perceive their own role, how they see other people's responsibilities, and how they develop strategies to coordinate the moments before, during, and after championship matches. Furthermore, we want to improve our comparison of the different games (CS:GO and R6) with the help of these methods.

The questionnaire developed for this article is still online, and we are receiving more answers that will allow us to update our analysis in the future.

Some of the questions raised by this article are also interesting for further exploration:

(1) To what extent does competition within eSport teams affect team performance? What factors increase competition and how do players, captains, and managers control them?

(2) How are eSport teams' coordination and communication capacities affected by the decisions taken during the formation of the teams?

(3) What collaborative necessities of eSport teams' players are not considered in the design of eSports' games? How do players adapt to the lack of such features?

Acknowledgments

We would like to thank all the players who answered the questionnaire. In particular, the eSport professionals who helped in the understanding of eSports — Alexandre Marinho and Lucas "Marduk" Rapinelli — and those who helped in the evaluation and divulgation of the questionnaire — Samuel "Sagaz" Byron, Willian "LuL4" Elias, Daniel "Danzera" Azevedo, Guilherme Barbosa, Diego Hads, Luigi Mataratzis, Ariane Parra, and Nicolle "Cherrygumms" Merhy. We would also like to thank CAPES for its financial support of this work.

References

Academy, U.S.S. 2017, May 11. Recognizing ESports as a Sport. Available from http://thesportjournal.org/article/recognizing-esports-as-a-sport/ [accessed 16 January 2018].

Beck, K. 2017a. Disney brings esports to TV in a major move. Available from http://mashable.com/2017/07/14/disney-esports/#Mfrb1nYoJPqm [accessed 5 January 2018].

Beck, K. 2017b. NBC Sports is getting into esports, starting with "Rocket League." Available from http://mashable.com/2017/06/21/rocket-league-nbc-television/#QeWHoJT1qOqa [accessed 5 January 2018].

Brown, B., and Bell, M. 2004. CSCW at Play: “There” As a Collaborative Virtual Environment. *In* Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work. ACM, New York, NY, USA. pp. 350–359. doi:10.1145/1031607.1031666.

Bullock, P. 2017, July 1. The Future of eSports on TV. Available from https://medium.com/the-cube/the-future-of-esports-on-tv-b2d1e77c58cc [accessed 5 January 2018].

Campion, M.A., Cheraskin, L., and Stevens, M.J. 1994. CAREER-RELATED ANTECEDENTS AND OUTCOMES OF JOB ROTATION. Academy of Management Journal **37**(6): 1518–1542. doi:10.2307/256797.

Cheung, V., Chang, Y.-L.B., and Scott, S.D. 2012. Communication channels and awareness cues in collocated collaborative time-critical gaming. *In* Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work. ACM. pp. 569–578.

Dafai, S.A. 2016. Conventions within eSports: Exploring Similarities in Design. : 16.

DeChurch, L.A., and Mesmer-Magnus, J.R. 2010. Measuring shared team mental models: A meta-analysis. Group Dynamics: Theory, Research, and Practice **14**(1): 1–14. doi:10.1037/a0017455.

Endsley, M.R. 1995. Toward a theory of situation awareness in dynamic systems. Human factors **37**(1): 32–64.

Eriksson, T., and Ortega, J. 2006. The adoption of job rotation: Testing the theories. ILR Review **59**(4): 653–666.

Freeman, G., and Wohn, D.Y. 2017a. eSports As An Emerging Research Context at CHI: Diverse Perspectives on Definitions. *In* Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems. ACM, New York, NY, USA. pp. 1601–1608. doi:10.1145/3027063.3053158.

Freeman, G., and Wohn, D.Y. 2017b. Understanding eSports Team Formation and Coordination. Computer Supported Cooperative Work (CSCW). doi:10.1007/s10606-017-9299-4.

Fuks, H., Raposo, A., and Gerosa, M.A. 2008a. The 3c collaboration model. *In* Encyclopedia of E-collaboration. IGI Global. pp. 637–644.

Fuks, H., Raposo, A., Gerosa, M.A., Pimentel, M., Filippo, D., and Lucena, C. 2008b. Inter-and intra-relationships between communication coordination and cooperation in the scope of the 3C Collaboration Model. *In* Computer Supported Cooperative Work in Design, 2008. CSCWD 2008. 12th International Conference on. IEEE. pp. 148–153.

Funk, D.C., Pizzo, A.D., and Baker, B.J. 2017. eSport management: Embracing eSport education and research opportunities. Sport Management Review. doi:10.1016/j.smr.2017.07.008.

Gergle, D., Kraut, R.E., and Fussell, S.R. 2004. Action as language in a shared visual space. *In* Proceedings of the 2004 ACM conference on Computer supported cooperative work. ACM. pp. 487–496.

Good, O.S. 2017, October 29. Olympic committee lays out expectations for esports’ inclusion. Available from https://www.polygon.com/2017/10/29/16566528/olympics-ioc-statement-esports [accessed 5 January 2018].

Graham, B.A. 2017, August 9. eSports could be medal event at 2024 Olympics, Paris bid team says. The Guardian. Available from http://www.theguardian.com/sport/2017/aug/09/esports-2024-olympics-medal-event-paris-bid-committee [accessed 5 January 2018].

Hallmann, K., and Giel, T. 2017. eSports – Competitive sports or recreational activity? Sport Management Review. doi:10.1016/j.smr.2017.07.011.

Hamari, J., and Sjöblom, M. 2017. What is eSports and why do people watch it? Internet Research **27**(2): 211–232. doi:10.1108/IntR-04-2016-0085.

Hamilton, W., Kerne, A., and Moeller, J. 2012a. Pen-in-hand Command: NUI for Real-time Strategy Esports. *In* CHI ’12 Extended Abstracts on Human Factors in Computing Systems. ACM, New York, NY, USA. pp. 1455–1456. doi:10.1145/2212776.2212483.

Hamilton, W., Kerne, A., and Robbins, T. 2012b. High-performance Pen + Touch Modality Interactions: A Real-time Strategy Game eSports Context. *In* Proceedings of the 25th

Annual ACM Symposium on User Interface Software and Technology. ACM, New York, NY, USA. pp. 309–318. doi:10.1145/2380116.2380156.

Heere, B. 2017. Embracing the sportification of society: Defining e-sports through a polymorphic view on sport. Sport Management Review. doi:10.1016/j.smr.2017.07.002.

Ikenami, R.K., Lipovaya, V., Sigette, E., Silva, E., and Duarte, F. 2018. The Ecosystem Approach as an Alternative to Navigate Across Innovation Environment. *In* Strategizing New Growth Avenues in an Evolving Global Context. São Paulo, Brazil.

Jenny, S.E., Manning, R.D., Keiper, M.C., and Olrich, T.W. 2017. Virtual(ly) Athletes: Where eSports Fit within the Definition of "Sport." Quest **69**(1): 1–18. doi:10.1080/00336297.2016.1144517.

Kaytoue, M., Silva, A., Cerf, L., Meira, W., Jr., and Raïssi, C. 2012. Watch Me Playing, I Am a Professional: A First Study on Video Game Live Streaming. *In* Proceedings of the 21st International Conference on World Wide Web. ACM, New York, NY, USA. pp. 1181–1188. doi:10.1145/2187980.2188259.

Kim, Y.J., Engel, D., Woolley, A.W., Lin, J.Y.-T., McArthur, N., and Malone, T.W. 2017. What Makes a Strong Team?: Using Collective Intelligence to Predict Team Performance in League of Legends. ACM Press. pp. 2316–2329. doi:10.1145/2998181.2998185.

Kow, Y.M., and Young, T. 2013. Media Technologies and Learning in the Starcraft Esport Community. *In* Proceedings of the 2013 Conference on Computer Supported Cooperative Work. ACM, New York, NY, USA. pp. 387–398. doi:10.1145/2441776.2441821.

Kozachuk, J., Foroughi, C.K., and Freeman, G. 2016. Exploring Electronic Sports: An Interdisciplinary Approach. Proceedings of the Human Factors and Ergonomics Society Annual Meeting **60**(1): 2118–2122. doi:10.1177/1541931213601479.

Lameiras, J., Almeida, P.L., and Garcia-Mas, A. 2014. Relationships between Cooperation and Goal Orientation among Male Professional and Semi-Professional Team Athletes. Perceptual and Motor Skills **119**(3): 851–860. doi:10.2466/25.PMS.119c32z4.

Leavitt, A., Keegan, B.C., and Clark, J. 2016. Ping to Win?: Non-Verbal Communication and Team Performance in Competitive Online Multiplayer Games. ACM Press. pp. 4337–4350. doi:10.1145/2858036.2858132.

Lipovaya, V., Ikenami, R.K., Rodrigues, G.M.S.M., and Duarte, F. 2017. Ecossistemas Inovadores eSports: o caso da empresa SUPERNOVA. *In* X Congresso de Administração, Sociedade e Inovação (CASI). Petrópolis, Brazil.

Livingston, I.J., Gutwin, C., Mandryk, R.L., and Birk, M. 2014. How players value their characters in world of warcraft. ACM Press. pp. 1333–1343. doi:10.1145/2531602.2531661.

Mason, W., and Clauset, A. 2013. Friends ftw! friendship and competition in halo: Reach. *In* Proceedings of the 2013 conference on Computer supported cooperative work. ACM. pp. 375–386.

Maxwell, J. 2009. The SAGE Handbook of Applied Social Research Methods. *In* 2nd edition. SAGE Publications, Inc., Thousand Oaks. doi:10.4135/9781483348858.

McClelland, P.J., Whitmell, S.J., and Scott, S.D. 2011. Investigating Communication and Social Practices in Real-time Strategy Games: Are In-game Tools Sufficient to Support the Overall Gaming Experience? *In* Proceedings of Graphics Interface 2011. Canadian Human-Computer Communications Society, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada. pp. 215–222. Available from http://dl.acm.org/citation.cfm?id=1992917.1992953 [accessed 15 January 2018].

Mintzberg, H. 1989. The structuring of organizations. *In* Readings in Strategic Management. Springer. pp. 322–352.

Morschheuser, B., Maedche, A., and Walter, D. 2017. Designing Cooperative Gamification: Conceptualization and Prototypical Implementation. ACM Press. pp. 2410–2421. doi:10.1145/2998181.2998272.

Newzoo. 2018. Key Resources | Games, Esports, Mobile | Market Size & Projections. Available from https://newzoo.com/resources/ [accessed 16 January 2018].

Senado Federal. 2018. Senado analisará regulamentação de esportes eletrônicos. Available from https://www12.senado.leg.br/noticias/materias/2017/12/26/senado-analisara-regulamentacao-de-esportes-eletronicos [accessed 16 January 2018].

Skubida, D. 2016. Can Some Computer Games Be a Sport?: Issues with Legitimization of eSport as a Sporting Activity. International Journal of Gaming and Computer-Mediated Simulations **8**(4): 38–52. doi:10.4018/IJGCMS.2016100103.

Taylor, T.L. 2012. Raising the stakes: e-sports and the professionalization of computer gaming. MIT Press, Cambridge, Mass.

Teruel, M.A., Navarro, E., González, P., López-Jaquero, V., and Montero, F. 2016. Applying thematic analysis to define an awareness interpretation for collaborative computer games. Information and Software Technology **74**: 17–44.

Toups, Z.O., Hammer, J., Hamilton, W.A., Jarrah, A., Graves, W., and Garretson, O. 2014. A framework for cooperative communication game mechanics from grounded theory. ACM Press. pp. 257–266. doi:10.1145/2658537.2658681.

Wagner, M.G. 2006. On the Scientific Relevance of eSports. *In* International Conference on Internet Computing. pp. 437–442.

Wilson, P. 1991. Computer supported cooperative work:: An introduction. Springer Science & Business Media.

Appendix

Questionnaire

(1) What is your age?
(2) What is your sex?
    a. Male
    b. Female
(3) What is the main game that you play?
(4) How many hours weekly, on average, you play this game?
(5) Do you work or study besides playing?
    a. Yes, both
    b. Yes, I work
    c. Yes, I Study
    d. No, eSports is my job
    e. Other (blank field to fill with answer)
(6) Is eSports your main source of income?
    a. Yes
    b. No
(7) Do your parents (or someone else) give you any financial support?
    a. Yes
    b. No
(8) What city are you from:
(9) Are you part of any team?
    a. Yes
    b. No

(10) If you want to, tell us the name of your team (We will not reveal this information, it will only be used to know you better)
(11) Have you ever participated in any eSports championship with prizes in money?
    a. Yes
    b. No
(12) Do you consider yourself a professional eSports player?
    a. Yes
    b. No
(13) Do a captain exist in your team (someone that leads the team during a match)?
    a. Yes, a captain that plays with the team during matches
    b. Yes, a captain that DOES NOT play with the team during matches
    c. Yes, some players rotate the role of captain
    d. No, this role of coach is divided among the team members
    e. No
    f. Other (blank field to fill with answer)
(14) If this role of captain exists, what do you expect of him moments before and during a match?
(15) Does your team have a coach?
    a. Yes, a coach that plays with the team during matches
    b. Yes, a coach that DOES NOT play with the team during matches
    c. No, this role of coach is divided among the team members
    d. No
    e. Other (blank field to fill with answer)
(16) If this role of coach exists, what do you expect of him moments before and during a match?
(17) Does your team have a manager that makes strategic decisions as which championships to compete, which marketing strategies to adopt, etc.?
    a. Yes, we have a manager that plays with the team
    b. Yes, we have a manager that DOES NOT play with the team
    c. No
(18) If your team has a manager, what do you expect of him moments before and during a match?
(19) Into what roles is your team divided during matches (e.g., in LoL there are shooters, assassins, tanks, fighters and supports)?
(20) And which of these roles is yours?
(21) Is there any rotation of these roles among players?
    a. Yes, we have already changed of role in our team between championships
    b. Yes, we have already changed of role in our team during a championship

    c. Yes, we have already changed of role in our team during a match
    d. No, we never changed roles

(22) What communication means does your team use during a match?
    a. Voice
    b. Video
    c. Text
    d. In-game visuals (e.g., ping, crouching our jumping with the avatar/character, etc.)
    e. Real-world visuals (e.g., signals, touches, etc.)
    f. Other (blank field to fill with answer)

(23) If you chose more than one, which is one does your team uses the most?

(24) Which tool(s) do you use to communicate during a match (e.g., Teamspeak, Discord, in-game mechanics, etc.)? Please, cite all of them, even if you do not use too much

(25) Tell us a bit about one or more situations in which you have to collaborate with someone of your team during a match (e.g., to destroy a structure, I usually tell the shooter of my team to help me attacking while I keep the distance with my mage)

(26) Before starting a championship match, what does your team use to do? Do you make some kind of concentration? How does it happen?

(27) What about after the match, what does your team do?

(28) Is there competition among the players of your own team?
    a. Yes
    b. No

(29) If this competition does exist, why does it happen? What motivates a player to seek a personal goal instead of a team goal? Tell us about some cases in which this competition occurs

(30) Do you think that is there any important question that we should have asked you about this topic? Which one would it be? How would you answer that question?

(31) Thanks for answering the questionnaire, leave any of your contact (e-mail, phone number, Facebook, Reddit, etc.) so that we can invite you to new research and share with you the material that our group produces about eSports! We will not fill your inbox with spam; we also do not like that.

Wilson A. and De Paoli, S. (2018): Reconsidering online reputation systems. In: Proceedings of 16th European Conference on Computer-Supported Cooperative Work - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2018_12

Reconsidering online reputation systems

Anna Wilson*†, and Stefano De Paoli*
*School of Social and Health Sciences, Abertay University, Dundee DD1 1HG, UK
†Faculty of Social Sciences, University of Stirling, Stirling FK9 4LA, UK
a.wilson@abertay.ac.uk, s.depaoli@abertay.ac.uk

Abstract. Social and socioeconomic interactions and transactions often require trust. In digital spaces, the main approach to facilitating trust has effectively been to try to reduce or even remove the need for it through the implementation of reputation systems. These generate metrics based on digital data such as ratings and reviews submitted by users, interaction histories, and so on, that are intended to label individuals as more or less reliable or trustworthy in a particular interaction context. We suggest that conventional approaches to the design of such systems are rooted in a capitalist, competitive paradigm, relying on methodological individualism, and that the reputation technologies themselves thus embody and enact this paradigm in whatever space they operate in. We question whether the politics, ethics and philosophy that contribute to this paradigm align with those of some of the contexts in which reputation systems are now being used, and suggest that alternative approaches to the establishment of trust and reputation in digital spaces need to be considered for alternative contexts.

Introduction

Trust is a fundamental component of social relations. It helps actors make decisions in situations where direct knowledge that can guide action and cooperation is not always immediately available. Trust helps reduce complexity in social interactions, allowing actors to take decisions in situations which entail some risk (Luhmann, 1979). Interactions in a digital environment are likely to require trust (Hsu et al., 2007; Usoro et al., 2007) even more than those in a

physical environment. While trust is often seen as a tri-partite relation between one individual (trustor) and another one (trustee) in relation to an object or outcome, it can also take a collective form in what is known as reputation, or how a community or group of people view the trustworthiness of another person or another entity.

The increase in the availability of digital data is having a significant impact on our opportunities to engage in social interactions and the ways in which they are enacted. Increased digitization leads to increased remote and mediated interactions. If we view humanity as a network: before the internet, interactions tended to be between nodes that were previously only separated by a few degrees; now the chances of creating a new connection/entering into a transaction or relationship with a previously very distant node are much higher, and the chances of the different parties to a social or economic transaction being physically co-located are much reduced (Shu and Chuang, 2011).

This has led to questions about how to establish trust in mediated interactions involving distant and/or unfamiliar actors, when:

- We don't know whether the person we're interacting/transacting with is who they say they are.
- We don't know whether they have the goods, skills or knowledge they claim to have.
- We don't know whether their digital presence will persist, and so whether we will have any continued relationship (and therefore a chance to reciprocate or for comeback).
- We can't rely on local knowledge and word-of-mouth (reputation).

One of the main ways in which online platforms have responded to this situation is through the development of reputation systems (Dellarocas, 2003; Jensen, Davis, & Farnham, 2002; Resnick et al., 2000). These are systems that collate data in the form of feedback, ratings, and digital interaction/transaction histories, process them through algorithms, and produce a synthetic and very often quantitative measure intended to give a guide to an individual's trustworthiness (Farmer and Glass, 2010).

In a context/mission creep mirroring that of other business-intelligence inspired data analytics (Wilson et al., 2017), such systems are becoming increasingly ubiquitous, no longer confined to the trading and expert knowledge-sharing sites they were originally developed for. Described by Masum and Zhang as a 'distributed court of opinion' (2004, n.p.) that will alleviate the strain on our overburdened 'individual processing capacity' (ibid.) in the face of vastly increased accessible data and so 'help the same number of hours in the day go further' (ibid.), great things are expected of them. It has been suggested they could play pivotal roles in the creation and maintenance of good governance,

transparency and accountability in public office and commerce, through either the creation of trust or even – paradoxically – the removal of the need for it (Litos and Zindros, 2017; Masum, Tovey and Newmark, 2012; Masum and Zhang, 2004; Picci, 2007). But as they, or components such as ratings systems, permeate into perhaps unexpected digital spaces – such as learning environments1, community support groups (see, e.g., http://supportgroups.com) or even online communities of criminals needing to trust each other in the exchange of services and goods such as hacking and botnets (Décary-Hétu and Dupont, 2013; Dupont et al., 2016) – and as China moves to introduce a mandatory social credit system that incorporates elements of online reputation systems (Botsman, 2017) that attempts to reduce individuals to single measures of quality, we need to ask questions about whether their design is commensurate with the intentions of the systems they are being brought into.

It is now fairly accepted in certain areas of research, that technology and technical artefacts (including information and communication technologies) are not politically or morally neutral. Winner (1980) argued that artefacts, very much like people, have their own politics which cause them to enact or contribute to particular types of ordered social system. He described the now well-known example of the low bridges on roads to Long Island from New York. Winner noticed that the low height of these bridges would exclude categories of people (those travelling on buses, generally working class people or African Americans) from certain actions, such as accessing a middle-class residential area. These low bridges thus embodied political decisions and enacted particular discriminations and exclusions.

In relation to reputation systems, an important question arises concerning the political and moral decisions that these systems embody and carry into the digital spaces they operate in. One episode of the TV series Black Mirror, Nosedive[2], takes the idea of ubiquitous reputation systems to the extreme; in so doing, it powerfully illustrates some of the political implications of reputation systems and their capacity to be the driver of social exclusion and inclusion. In it, people use an app on their mobile phone to rate each other during or after any real interaction. In a plot move that has echoes of the developing Chinese social credit system (Botsman, 2017; Hvistendahl, 2017), those with high scores have access to better apartments and other perks. On the other hand, those with low scores become social outcasts. This reputation system, then, constitutes an instrument for both upward and downward social mobility. Similarly to the low bridges

1 Virtual learning environments and learning management systems are increasingly making use of both the conceptual models and user-interfaces of business-intelligence applications, with examples such as the popular CANVAS system using 3-star scales to indicate student performance.

2 https://www.theatlantic.com/entertainment/archive/2016/10/black-mirror-nosedive-review-season-three-netflix/504668/

described by Winner, the reputation system in Nosedive embodies a politics with rules of exclusion and inclusion which are enacted through and by the artefact.

We suggest that conventional reputation systems are loaded with not just the values they are designed for (trust, honest behaviour, reliability), but also a more extended and subtle value-system: the political and ethical paradigm of the competitive, capitalist free market based on self-interested individuals. Such systems are underpinned by a view of reputation that implicitly (and sometimes even explicitly) commoditizes it, positioning it as a capital (most explicitly in work such as Gandini (2016)) that is inherent to individuals, who can accumulate it, lose it and occasionally even speculate on it.

This might be appropriate for a digital system that is intended to serve as a competitive market, for example an e-commerce website, or to function within a platform capitalist model. However, this may not be the case in other contexts, where a different political, ethical or philosophical paradigm underpins the construction or enactment of the digital space.

In the following, we describe the main features of conventional reputation systems and show why we believe they embody and enact a fundamentally market-based, capitalist paradigm. We then examine various contexts in which such systems, or parts of them, operate, including trading sites (eBay/Etsy/gig economic sites), expert question-and-answer (Q&A), and supportive discussion forums, and ask whether the properties and features of these systems are likely to encourage the kinds of behaviours that participants in and designers of these sites may wish for. Finally, we offer some preliminary observations associated with a project we are working on in the area of Collective Awareness Platforms for Social Innovation, whose goal is to facilitate a novel form of welfare – termed commonfare (Fumagalli and Lucarelli, 2015) – among people who have experienced conditions of poverty or precariousness. We argue that the dominant model of a reputation system would clash with what the project aims to achieve.

Trust and reputation systems in digital spaces

Common features of online reputation systems

Online reputation systems are systems that draw on data about a user's activities to generate an indication of that user's standing within one or more online communities (Dellarocas, 2003; Jensen, Davis, and Farnham, 2002; Resnick et al., 2000). In some ways similar to the points systems and leader-boards common to

online games, in which points are sought competitively and assigned by the game; the "capital" nature of such points is made clear in those games that allow players to "spend" their points within the game-world.

Reputation systems outside of games have a stronger focus on providing users with a metric on which to base judgments about whether to trust other users or select them as partners for a transaction. They are now default parts of the design of e-commerce sites, where items are bought and sold in conventional financial transactions. They are also integral to the increasing number of sites based on a "gig" (Friedman, 2014) or "sharing" (Hamari, Sjöklint, and Ukkonen, 2015) economic model. (In the former, members offer their skills and services for money but in a freelance capacity; in the latter, they provide or/and seek resources such as tools, transport or accommodation without the exchange of money.) In addition, many expert Q&A sites (usually based on discussion forum rather than trading structures) employ reputation systems so that questioners can judge whether or not to trust an answer, or community members can build up their own reputation as experts (see, e.g., Movshovitz-Attias et al., 2013). For participants in these latter sites, high reputation scores may also be seen as badges of achievement or honour – measures of kudos, as indicated by the name of the reputation scores in the online expert coder community StackOverflow (Movshovitz-Attias et al., 2013; Bosu et al., 2013). The inclusion of reputation systems in a digital space may thus also be seen as a form of gamification, providing motivation to contribute more and higher quality postings or items in a knowledge-sharing community.

Reputation systems can base reputation measures on data from a range of sources, processed in a range of ways (Costagliola, Fuccella and Pascuccio, 2014; Hendrikx, Bubendorfer and Chard, 2015; Vavilis, Petković and Zannone, 2014). They may employ data generated directly from a user's activities, such as how many times they visit a site, how long they spend on a site, how many transactions they complete, the ratio of completed to started transactions, how many contributions they make to a discussion, how many network ties they have, and so on. They may also draw on ratings of that user's contributions/behaviour provided by other users: for example, through "likes," up- and down-votes, ratings against particular reputation-items such as helpfulness, reliability, promptness etc., or qualitative feedback in the form of text-based reviews. When reputation systems are intended to support transactions of a trading nature (whether as part of the conventional, gig or sharing economy), an entity's reputation score might be based on customer feedback about reliability, product quality, speed of response, etc. When they are intended to support expert discussion forums or interest groups, reputation scores may be based on other users' judgments of the quality of an individual's contributions to the site, number of contributions, and so on. In either case, reputation metrics are intended to serve as proxies (Floridi, 2015) for

prior experience and personal knowledge, on the basis of which predictions of future interactions can be made.

Whichever factors are included in a reputation system, they are often used to generate a numerical measure of the user's overall behaviour/reputation/ranking within the relevant community (despite Masum and Zhang's caution that 'No person can be reduced to a single measure of "quality"' (2004, np)). Reputation "scores" may be made public to other community members, so that they can make decisions about how and with whom they interact; or they may be known only to the site administrators (or an automated process) and used to make decisions about allowing or removing privileges within, and even access to, services and users within the space. In the former case, they will also be visualized on the interface of the service (e.g., using star-ratings or badges). Scores may be aggregates or averages; the data used to calculate these scores may be unweighted or weighted according to a range of factors, including the reputation of the user submitting the ratings and the age of the rating.

Trust and reputation as forms of capital

The notion of online reputation has received substantial attention in recent years, with some authors suggesting that the increasing digitization of transactions and interactions is leading to a "reputation society" (Masum, Tovey and Zhang, 2012) and others proposing that reputation is in fact social capital in a "reputation economy" (Gandini, 2016).

As evident from the description in the previous sub-section, online reputation systems have been developed for two general purposes: as tools to help users of web-based platforms make decisions about whom to trust; and as motivators for more and higher quality participation in certain web-based activities or communities. Such systems are based on the premise that 'reputation becomes visible, tangible and, under certain conditions, even measurable … through algorithms and metrics that elaborate online reputation scores' (Gandini, 2016, 28). Some authors suggest that this kind of measurement and sharing of reputation information could radically shift the balance of power in society, as 'peer networks will confer legitimacy on people emerging from the grassroots' (Newmark, 2012, ix).

We can explore more what kind of politics may be embodied in conventional reputation system designs, and see how this politics is re-inscribed back into online communities. These systems appear to be based on individualism, the free market as the ideal (political) economy and liberalism as the essence of social relations. For example, Dellarocas, one of the most influential theorists of these type of systems, states that '[t]he new platforms may be all about harnessing crowds and communities, but in the end, those crowds and communities are

nothing but a sum of individuals' (2010, 33), an attitude that neglects the sometimes strong and complex social, political and cultural mechanisms that couple individuals and result in emergent, collective behaviour. Similarly, Picci explicitly positions his arguments as rooted in rational choice theory, positing that 'individual social actors act to advance their self-interest' (2007, np) and claiming that reputation systems 'allow selection forces to weed out the least fit' (ibid.). Gandini's claim that reputation is social capital rests on the belief that it is 'an eminently economic concept' (2016, 30) that 'functions as a form of currency enabling trust among strangers' (ibid., 32) and that is 'a resource that may be mobilized and that remains with the individual … as a capital that is invested, traded or managed … as an investment in social relations with expected economic return' (ibid., 36), a view that combines individualism with a clear commoditization of reputation.

One might ask whether reputation systems as currently developed are more likely to reinforce self-interested individualism, since they are grounded in a methodological individualism which sees social groups as aggregations of individuals, each aiming at self-satisficing egoistic behaviour, under the often not explicit idea that this is done for the benefit of the whole group. As Adam Smith famously stated, 'It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own self-interest' (Smith 1838, 7).

Paradoxically, some aspects of this kind of thinking appear to have been taken to extremes in the (nominally socialist) Chinese government's recent experiments with and planned national roll-out of a combined social and personal-financial credit system (Botsman, 2017; Hvistendahl, 2017). In these, conventional measures of financial credit-worthiness are being combined with behavioral metrics including shopping habits, friendship networks and the sharing of 'positive energy' (Botsman, 2017, n.p.) online to produce a single trustworthiness or social credit score.

However, if reputation is viewed as a currency or marketable commodity, resulting from action of self-interested individuals, then it may be exposed to the same risks and problems that arise in financial markets, including questions related to ownership, fairness and control. Indeed, the global financial crisis has led to renewed questions as to the validity of competitive, free-market models and suggestions that approaches that recognize the strong coupling of different components in the system should be developed (Helbing and Kirman, 2013). Within the economics of reputation and trust that reputation systems are helping to create, there is already evidence for the kind of problems that arise when financial gain can be made by adopting certain behaviours, including the use of multiple or fake personas to acquire undeserved reputational credit/value, exaggerated reciprocity, individualised reciprocity resulting in clique formation, retaliation and clique-based attacks.

For example, there have been several studies of eBay's reputation system and the impact it has on participation in the system (see, for example, Cabral and Hortaçsu, 2010; Dellarocas, Fan and Wood, 2004; Houser and Wooders, 2006; Hui et al., 2014; Resnick et al., 2000; Resnick and Zeckhauser, 2002; Resnick, Zeckhauser and Swanson, 2006). The main findings of this research suggest that feedback contributions on eBay are not strongly driven by altruism (Dellarocas, Fan and Wood, 2004), and instead are more strongly driven by an expectation of reciprocity. Resnick et al. (2000) suggest that users not only reciprocate but also retaliate. They also suggest that users of the site become less likely to participate in the feedback process once they have accumulated experience (and "respectable" reputation scores). This observation is consistent with the suggestion that users' participation in the feedback process is not strongly driven by altruism, as it may imply that once users have built up a secure reputation, they no longer feel the need to elicit ratings from others by providing ratings themselves. Resnick, Zeckhauser and Swanson (2006) showed that reputation is, however, important, and that the same items, sold by the same seller under two different identities, attracted an 8% lower price when sold through a newly-established identity with low reputation, as compared to the seller's "real" (well-established, high reputation) identity. Cabral and Hortaçsu (2010) studied the impact of negative feedback, finding that the first time a seller receives negative ratings/feedback has a more significant impact on his/her sales than subsequent negative ratings, but also that once a seller receives a negative rating, they are much more likely to receive more. They also found that sellers with low reputations are more likely to exit the system. Reputation systems, then, may not only enact a market-based, accumulative and acquisitive capitalist paradigm in whatever digital spaces they are employed – but also risk introducing behaviours that are detrimental to the health and sustainability of those spaces.

While a market-based view of reputation may be acceptable and even desirable in a business-focused trading site, it may undermine the intentions and purported values of other types of site. For example, although expert Q&A sites could be viewed as markets for knowledge, with competition among providers to supply the best quality knowledge, they are not real markets in the sense that there is no obvious cost to those seeking (and presumably consuming) the knowledge on offer and knowledge-providers retain the knowledge that they give out. Instead, a closer comparison might be with school or university learning environments, or sites of professional learning, where knowledge, once created, can be distributed and shared at no loss to any party to the sharing transaction. Rather than the power dynamics of a market, governed by competition and differentiation in wealth, expert Q&A sites are more likely to be characterised by dynamics of pride and commitment to the advancement and promotion of particular forms of knowledge and skill. In this kind of context, reputation might still take the form

of something to be accumulated, but not so much at the expense of other actors competing for the same resources and capital.

Indeed, it seems that some reputation systems used in expert Q&A sites, such as that incorporated in StackOverflow, reflect some of these differences. StackOverflow is a Q&A site where programmers can ask and answer questions relating to technical issues, and it has probably the best-known and most elaborately-developed reputation system in a Q&A site (Bosu et al., 2013; Hart & Sarma, 2014; Movshovitz-Attias et al., 2013). In StackOverflow's reputation system, users can up-vote and down-vote questions and answers provided by others, actions that not only contribute to reputation-building but also move questions up and down in terms of the order of display, and so make them more or less visible. Users gain and lose reputation in a variety of ways, including through the up- and down-voting of questions; there are many more ways to gain reputation than to lose it. The most significant way to lose reputation points involuntarily is if a post is flagged as offensive or spam; points can also be "spent" (transferred to another user) in a bounty system for those seeking quick and accurate answers to complex or esoteric questions.

In StackOverflow, points are converted into privileges: for example once a user has 15 points, they can vote up a question or answer; once they have, 20, they can talk in a chat; once they have 125, they can vote down questions or answers; and so on. At 1500 points users are allowed to add new tags to the site (questions are tagged as corresponding to particular topic areas, such as SQL or java); at 200, users can edit other users' questions and answers. At 10000 points users gain moderation rights; at 25000, they have access to the site's analytics. Thus there are incentives to build one's reputation that go beyond the acquisition of reputation for its own sake, or in order to gain the trust of other users.

However, this reputation system is still grounded in an individualistic, accumulative and competitive paradigm, which may have negative consequences for the diffusion of professional knowledge. For example, Movshovitz-Attias et al. (2013) found that while the majority of questions on the site were posted by novice users with low reputations, on average higher reputation users ask more questions than lower reputation users, simply because they contribute more often to the site. StackOverflow has also been found to (unintentionally) exclude or discourage female participants (Vasilescu et al., 2012), which has been partially attributed to the reward system. Thus StackOverflow's reputation system, while already incorporating some features that better reflect the aims of expert-community knowledge sharing and creation, may still to some extent undermine its aims and ethos.

Moving away from the traditional spaces in which reputation systems were developed, systems based on the same principles are also increasingly being incorporated into digital spaces that set out to bypass commercial transactions and achieve cooperative or mutualistic transactions. For example, the

accommodation-arranging platform Couchsurfing.com positions itself as setting out to achieve a social good: 'We envision a world made better by travel and travel made richer by connection. Couchsurfers share their lives with the people they encounter, fostering cultural exchange and mutual respect' (Couchsurfing, 2016). Couchsurfing.com relies on substantial levels of trust between strangers, as users share their homes with each other without any monetary exchange.

Lauterbach (2009) showed that there are significant levels of both direct and generalized reciprocity within the overall couchsurfing community. Couchsurfing's reputation system is based on systems used in conventional economic trading sites but has two unusual features. The first is in its use of friendship ties. Users can identify the type of relationship they have with other users, choosing from: Haven't met yet, Acquaintance, CouchSurfing friend, Friend, Good friend, Close friend, and Best friend. Couchsurfers who have hosted or stayed with other members are permitted to submit private feedback (to Couchsurfing) and public references for 14 days after a stay. Members must have a couch request with the "Yes" "Maybe" or "Confirmed" status in order to leave a Surf/Host reference. Other members may create references under the "Other" or "Friend" reference designations (as opposed to "Surf" or "Host"). Users' publicly visible reputation information is simply the number of references they have been given, and the number of those that are positive and have been confirmed (i.e., the user has confirmed the host/guest exchange). Other users can see free text references left by former guests/hosts.

It seems that this qualification of feedback based on the nature of relationships may be an attempt to mitigate the pure free-market nature of a conventional ratings-based system, in which every opinion counts the same, no matter how well-informed. However, Couchsurfing has a second unusual feature, which may be an example of how a reputation system can undermine the stated ethos of a platform. After some years of operating with the system described above, Couchsurfing.com introduced an additional "vouching" system, to allow some users to increase their reputation levels. This very restrictive system allows users to vouch for other users only if they have received three or more "vouches" themselves, effectively restricting vouching to an elite core: in, 2009, only 6.8% of members were able to vouch (Lauterbach et al., 2009). Thus the use of a conventional reputation system – albeit with some modifications – may in fact represent a misalignment with Couchsurfing's stated values of opening up sociocultural spaces and recognizing the contribution to this endeavour made by anyone who is willing to open up their home to a stranger.

Reputation systems in supportive groups

Finally, we consider another context in which reputation systems sometimes appear: that of supportive discussion forum sites. On the surface, these may seem to be similar to the expert Q&A forums considered in the previous section – discussion boards to which users can post questions that they are seeking answers to from community members with similar interests and pre-occupations. However, there are some fundamental differences to the aims and use of such sites.

First, expert Q&A sites such as StackOverflow are professional/technical interest community sites. Their users tend to be people who already have some degree of technical expertise (and therefore knowledge and cultural capital) and are seeking more. Several things follow from this:

- Questions on sites such as StackOverflow are technical in nature, seeking specific solutions to specific coding, implementation or operating system problems.
- They are likely to have answers which can be clearly judged as right, wrong or useful, depending on whether these answers lead to solutions that the questioner (and other members of the community) can implement. Where there may be more than one correct answer/workable solution, some will be more efficient or simpler to implement than others, and can be judged better on those grounds.
- Because users have some existing level of expertise, their judgment as to the value of answers might be expected to be reasonably reliable.
- Users are often enthusiasts for their work, and so are discussing something they enjoy doing. They are also proud of their expertise and are keen to provide answers if they have them.
- Questions (and answers) on sites such as StackOverflow are almost never personal or emotional; they are rarely likely to be of dramatic importance to the questioner's life or living conditions.

In contrast, the stories that may be told, and the advice and guidance sought and given on community support discussion forums, for example relating to health issues or financial problems, may relate to issues which are of substantial personal significance to users. There are many such communities, some facilitated by charities, health systems, or other authoritative figures or structures, but others having a more grass roots or community-driven character (see, for example, Barak, Boniel-Nissim and Suler, 2008; Chung, 2013). Many are associated with particular illnesses, whether physical or mental (see, for example, Eysenbach et al., 2004; Griffiths et al., 2009; Wright and Bell, 2003). Some discussion forums and mailing lists have developed to provide a safe space for minority groups such as the LGBTQI community (Mehra, Merkel, and Bishop, 2004). Others provide

discussion forums for larger groups, a good example being the parenting forum Mumsnet (Pedersen and Smithson, 2013).

While reputation scores are not yet widespread among such sites, they have found their way into some of them. For example, the set of discussion forums hosted by the platform supportgroups.com, which includes forums dedicated to financial problems, homelessness, anxiety, and other mental and physical health issues, has a linked reputation system so that users can acquire points for contributing across the forums they contribute to. The use of reputation systems in digital spaces that might attract vulnerable, socially-isolated or excluded people may be of particular concern. There is a well-established correlation between poverty/financial uncertainty and depression/anxiety (see, for example, Belle Doucet, 2003; Galea et al., 2007; Murali and Oyebode, 2004; Murphy and Athanasou, 1999; Paul and Moser, 2009; Vinokur, Price and Kaplan, 1996), which is not surprising given the potential for experiences of precariousness, social exclusion and social isolation, and feelings of inadequacy and decreasing hope. While people may well have positive stories and strategies to share, they may often be describing how they overcame a difficulty that was quite an unpleasant experience. Similarly, those visiting the site in order to find help and advice may well be seeking the emotional, as well as practical, support that can be provided by a community of people experiencing similar difficulties. We might speculate on the potential impact of inscribing a capitalist-oriented reputation system into such an environment. While on the one hand users might value trust creation processes as they decide who to interact with and seek support from, it is easy to imagine situations in which reputation scores might have negative impacts, for example on users' self-esteem. Given the value-system inherent in the design of conventional reputation systems, reputation may represent another form of capital in which users can find themselves to be poor, and so another benchmark of failure, inadequacy and inequity.

Toward the Commonfare.net platform: Trust, reputation and shared values

We are involved in a Collective Awareness Platform for Social Innovation project which is currently building a mobile and web platform called commonfare.net. The goal of this digital space is to foster a new form of bottom-up, community welfare, called commonfare (Fumagalli and Lucarelli, 2015). The project hopes to help alleviate the consequences of the 2008 global financial crisis in Europe and the failure of state-based approaches to welfare to improve the living conditions of those at risk of or experiencing precarity and social exclusion. The project (Botto and Teli, 2017) adopts a participatory design approach and the

original intention was to include a fairly conventional reputation system based on user ratings. Evidence from both the analysis of existing reputation systems and preliminary results from working with community members contributing to the design process strongly suggested that this might undermine the values and ethos of the site. Moreover it became clear that having solutions promoting this ethos is far more important to them than having access to individualistic reputation metrics.

In the following, we provide a brief description of the commonfare concept and of the Commonfare.net platform and initially suggest that in this platform genuine trust is unlikely to be facilitated by conventional reputation systems based on an individualistic, acquisitive market paradigm. We then present a set of points which will guide our future work designing a reputation system for the platform.

The Commonfare.net project is dedicated to the development of a mobile-first, web-based platform through which to improve the lives of people experiencing economic and social exclusion or precarity in Europe, through the promotion and facilitation of commonfare, an alternative approach to social welfare (Fumagalli and Lucarelli, 2015). A commonfare approach is grounded in the recognition that the social and economic are not separate spheres, but instead are inextricably and intricately connected. Commonfare is:

- bottom-up
- socially equitable
- cooperative

Key features of a commonfare approach include proper management of the common (both physical commons such as water, land and so on and immaterial commons such as knowledge and affect); an unconditional, basic income for all; and the development of complementary financial circuits. Any digital space that attempts to encourage a commonfare approach must therefore have design processes and features that are consistent with its core principles of bottom-up, socially equitable and cooperative action.

Participants in the design process include individuals and community members representing unemployed and precarious young people (Croatia), precarious workers (Italy and the Netherlands), non-Western migrants (the Netherlands) and benefit recipients (the Netherlands). Commonfare.net is intended to be a collective awareness platform that facilitates the development of commonfare approaches to social welfare among its users. Commonfare.net will offer a complementary channel for the provision of social welfare, allowing users to take better advantage of State offerings as well as to create their own alternative support and empowerment mechanisms. (For more detail, see Botto and Teli (2017).)

As is evident from the above, the philosophy behind commonfare.net is one that values the provision of mutual support and activities that lead to communal benefit rather than self-interested individualism. To an extent, it is more similar to

the case of community support groups than to those of an e-commerce website or a technical Q&A site. This means that cooperation and collaborative action will be essential to the development of a strong and valuable commonfare. To achieve this, the platform will offer various ways for people to interact; commonfare.net will be a digital space that enables information provision, inspiration/motivation and community building through story-telling, and cooperative/mutually beneficial actions. This last might include exchanges of goods, skills, knowledge or services in a sharing economy, group creation/cooperation, and forums for supportive Q&A. Trust will be important in facilitating and encouraging all of these interactions and some kind of trust facilitation or reputation system is needed; however our analysis of existing reputation systems and initial results from empirical research show the need for a novel approach, which is not based on individualistic principles. Rather, what is needed is an approach that reinforces relationality and community cooperation. Our preliminary proposals for the reputation system for the commonfare.net platform are that it should be based on the following explicitly political, values-driven principles:

- Rejection of individualism in the face of a widespread desire to feel part of a community with shared values, especially one which cooperates and acts in a mutualistic way to increase the quality of life of the many.
- Valuing self-determination, autonomy, and freedom from conditionalities such as those imposed by the State, non-governmental authorities such as NGOs, and capitalist entities such as employers and big businesses, when they provide welfare support and help. The reproduction of trust models often associated with these entities will not necessarily facilitate the achievements of the project goals.
- Acknowledgement that building (and warranting) trust in the platform will be as or even more important than trust in the individuals a user might potentially interact with.
- Recognition of the danger of creating, and a desire to avoid, new forms of rather fragmented solidarity that may result in overly-segregated group formation and hence obstacles to the diffusion of knowledge and good practice. This may necessitate the public availability and active sharing of information that shows the different groups' levels of contribution and commitment to the shared goal of building the commonfare.

These principles, which have emerged strongly and consistently in our participatory research activities, appear to be fundamentally at odds with the methodological individualism of a competitive, acquisitive market in reputation as a form of capital. They thus direct us to re-think reputation and trust facilitation in our developing digital space.

Acknowledgments

This work is supported by the European Commission under the H2020-ICT-2015 Programme (grant nr.: 687922).

References

Barak, A., Boniel-Nissim, M., and Suler, J. (2008): 'Fostering empowerment in online support groups', *Computers in Human Behavior*, vol. 24, no. 5, September 2008, pp. 1867-1883.

Belle Doucet, D. (2003): 'Poverty, inequality, and discrimination as sources of depression among US women', *Psychology of Women Quarterly*, vol. 27, no. 2, June 2003, pp. 101-113.

Bosu, A., Corley, C. S., Heaton, D., Chatterji, D., Carver, J. C., and Kraft, N. A. (2013): 'Building reputation in stackoverflow: an empirical investigation', in *Proceedings of the 10th Working Conference on Mining Software Repositories*, IEEE Press, pp. 89-92.

Botsman, R. (2017): 'Big data meets Big Brother as China moves to rate its citizens', *Wired, 21/10/2017*. Retrieved 9th of February 2018 from http://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion.

Botto, F. and Teli, M. (2017): 'PIE News. A public design project toward commonfare', *The Journal of Community Informatics*, vol. 13, no. 2, no date, pp. 87-105.

Cabral, L., and Hortacsu, A. (2010): 'The dynamics of seller reputation: Evidence from eBay', *The Journal of Industrial Economics*, vol. 58, no. 1, March 2010, pp. 54-78.

Chung, J. E. (2013): 'Social interaction in online support groups: Preference for online social interaction over offline social interaction', *Computers in Human Behavior*, vol. 29, no. 4, July 2013, pp. 1408-1414.

Costagliola, G., Fuccella, V., and Pascuccio, F. A. (2014): 'Towards a trust, reputation and recommendation meta model', *Journal of Visual Languages & Computing*, vol. 25, no. 6, December 2014, pp. 850-857.

Couchsurfing (2016). *About us*. Retrieved 9th of February 2018 from http://www.couchsurfing.com/about/about-us/.

Décary-Hétu, D., and Dupont, B. (2013): 'Reputation in a dark network of online criminals', *Global Crime*, vol. 14, no. 2-3, May 2013, pp. 175-196.

Dellarocas, C. (2003): 'The digitization of word of mouth: Promise and challenges of online feedback mechanisms', *Management science*, vol. 49, no. 10, October 2003, pp. 1407-1424.

Dellarocas, C., Fan, M., and Wood, C. A. (2004): 'Self-interest, reciprocity, and participation in online reputation systems', *MIT Sloan Working Papers No. 4500-04*. Retrieved 9th of February from https://ssrn.com/abstract=585402.

Dupont, B., Côté, A. M., Savine, C., and Décary-Hétu, D. (2016): 'The ecology of trust among hackers', *Global Crime*, vol. 17, no. 2, March 2016, pp. 129-151.

Eysenbach, G., Powell, J., Englesakis, M., Rizo, C., and Stern, A. (2004): 'Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions', *BMJ*, vol. 328, no. 7449, May 2004, pp. 1166.

Floridi, L. (2015): 'A proxy culture', *Philosophy & Technology*, vol. 28, no. 4, October 2015, pp. 487-490.

Farmer, R., and Glass, B. (2010): *Building web reputation systems*, O'Reilly Media, Inc.

Friedman, G. (2014): 'Workers without employers: shadow corporations and the rise of the gig economy', *Review of Keynesian Economics*, vol. 2, no. 2, April 2014, pp. 171-188.

Fumagalli, A., and Lucarelli, S. (2015): 'Finance, Austerity and Commonfare', *Theory, Culture & Society*, vol. 32, no. 7-8, October 2015, pp. 51-65.

Galea, S., Ahern, J., Nandi, A., Tracy, M., Beard, J., and Vlahov, D. (2007): 'Urban neighbourhood poverty and the incidence of depression in a population-based cohort study', *Annals of Epidemiology*, vol. 17, no. 3, no date, pp. 171-179.

Gandini, A. (2016): *The reputation economy: Understanding knowledge work in digital society*. Springer.

Griffiths, K. M., Calear, A. L., Banfield, M. A., and Tam, A. (2009): 'Systematic review on Internet Support Groups (ISGs) and depression (2): What is known about depression ISGs?', *Journal of medical Internet research*, vol. 11, no. 3, no date, e41.

Hamari, J., Sjöklint, M., and Ukkonen, A. (2015): 'The sharing economy: Why people participate in collaborative consumption', *Journal of the Association for Information Science and Technology*, vol. 67, no. 9, June 2015, pp. 2047-2059.

Hart, K., and Sarma, A. (2014, June): 'Perceptions of answer quality in an online technical question and answer forum', in *Proceedings of the 7th International Workshop on Cooperative and Human Aspects of Software Engineering*, ACM, pp. 103-106.

Helbing, D., and Kirman, A. (2013): 'Rethinking economics using complexity theory'. Retrieved 8th of February 2018 from http://www.soms.ethz.ch/paper_economics_complexity_theory

Hendrikx, F., Bubendorfer, K., and Chard, R. (2015): 'Reputation systems: A survey and taxonomy', *Journal of Parallel and Distributed Computing*, vol. 75, January 2015, pp. 184-197.

Houser, D., and Wooders, J. (2006): 'Reputation in auctions: Theory, and evidence from eBay', *Journal of Economics & Management Strategy*, vol. 15, no.2, April 2006, pp. 353-369.

Hsu, M. H., Ju, T. L., Yen, C. H., and Chang, C. M. (2007): 'Knowledge sharing behavior in virtual communities: The relationship between trust, self-efficacy, and outcome expectations', *International journal of human-computer studies*, vol. 65, no. 2, February 2007, pp. 153-169.

Hui, X., Saeedi, M., Shen, Z., and Sundaresan, N. (2014): 'From lemon markets to managed markets: the evolution of ebay's reputation system', *Working Paper, Ohio State University*.

Hvistendahl, M. (2017): 'Inside China's vast new experiment in social ranking', *Wired, 12/14/2017*. Retrieved 8th of February from https://www.wired.com/story/age-of-social-credit/.

Jensen, C., Davis, J., and Farnham, S. (2002): 'Finding others online: reputation systems for social online spaces', in *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, pp. 447-454.

Lauterbach, D., Truong, H., Shah, T., and Adamic, L. (2009, August): 'Surfing a web of trust: Reputation and reciprocity on couchsurfing.com', in *CSE'09: Proceedings of the International Conference on Computational Science and Engineering*, IEEE, pp. 346-353.

Litos, O. S. T., and Zindros, D. (2017): 'Trust is risk: a decentralized financial trust platform', in *International Conference on Financial Cryptography and Data Security*, Springer, Cham, pp. 340-356.

Luhmann, N. (1979): *Trust and Power: Two Works by Niklas Luhman*. Chichester: John Wiley and Sons.

Resnick, P., Kuwabara, K., Zeckhauser, R., and Friedman, E. (2000): 'Reputation systems', *Communications of the ACM*, vol. 43, no. 12, December 2008, pp. 45-48.

Masum, H., and Zhang, Y. C. (2004): 'Manifesto for the reputation society', *First Monday*, vol. 9, no. 7, July 2004, http://firstmonday.org/ojs/index.php/fm/article/view/1158

Masum, H., Tovey, M., and Newmark, C. (2012): *The reputation society: How online opinions are reshaping the offline world*. Cambridge, MA: MIT Press.

Mehra, B., Merkel, C., and Bishop, A. P. (2004): 'The internet for empowerment of minority and marginalized users', *New media & society*, vol. 6, no. 6, December 2004, pp. 781-802.

Movshovitz-Attias, D., Movshovitz-Attias, Y., Steenkiste, P., and Faloutsos, C. (2013, August): 'Analysis of the reputation system and user contributions on a question answering website: Stackoverflow', in *ASONAM: Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, IEEE, pp. 886-893.

Murali, V., and Oyebode, F. (2004): 'Poverty, social inequality and mental health', *Advances in Psychiatric Treatment*, vol. 10, no. 3, May 2004, pp. 216-224.

Murphy, G. C., and Athanasou, J. A. (1999): 'The effect of unemployment on mental health', *Journal of Occupational and Organizational Psychology*, vol. 72, no. 1, March 1999, pp. 83-99.

Newmark, C. (2012): 'Foreword: Trust, Reputation Systems, and the Immune System of Democracy', in *The Reputation Society: How Online Opinions Are Reshaping the Offline World*. Cambridge, MA: MIT Press.

Paul, K. I., and Moser, K. (2009): 'Unemployment impairs mental health: Meta-analyses', *Journal of Vocational Behavior*, vol. 74, no. 3, June 2009, pp. 264-282.

Pedersen, S., and Smithson, J. (2013): 'Mothers with attitude—How the Mumsnet parenting forum offers space for new forms of femininity to emerge online', *Women's Studies International Forum*, vol. 38, May-June 2013, pp. 97-106.

Picci, L. (2007): 'Reputation-based governance', *First Monday*, vol. 12, no. 9, September 2007, http://firstmonday.org/ojs/index.php/fm/article/view/2010/1885

Resnick, P., and Zeckhauser, R. (2002): 'Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system' in Michael R. Baye (ed.) *The Economics of the Internet and E-commerce* (Advances in Applied Microeconomics, Volume 11) Emerald Group Publishing Limited, pp.127-157

Resnick, P., Zeckhauser, R., Swanson, J., and Lockwood, K. (2006): 'The value of reputation on eBay: A controlled experiment', *Experimental economics*, vol. 9, no. 2, June 2006, pp. 79-101.

Shu, W., and Chuang, Y. H. (2011): 'The perceived benefits of six-degree-separation social networks', *Internet Research*, vol. 21, no. 1, pp. 26-45.

Smith, A. (1838): *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: Black and Tait.

Usoro, A., Sharratt, M. W., Tsui, E., and Shekhar, S. (2007): 'Trust as an antecedent to knowledge sharing in virtual communities of practice', *Knowledge Management Research & Practice*, vol. 5, no. 3, August 2007, pp. 199-212.

Vasilescu, B., Capiluppi, A., and Serebrenik, A. (2012, December): 'Gender, representation and online participation: A quantitative study of stackoverflow', in *Proceedings of the International Conference on Social Informatics*, IEEE, pp. 332-338.

Vavilis, S., Petković, M., and Zannone, N. (2014): 'A reference model for reputation systems', *Decision Support Systems*, vol. 61, May 2014, pp. 147-154.

Vinokur, A. D., Price, R. H., and Caplan, R. D. (1996): 'Hard times and hurtful partners: How financial strain affects depression and relationship satisfaction of unemployed persons

and their spouses', *Journal of Personality and Social Psychology*, vol. 71, no. 1, July 1996, pp. 166-179.

Wilson, A., Thompson, T. L., Watson, C., Drew, V., and Doyle, S. (2017): 'Big data and learning analytics: Singular or plural?', *First Monday*, vol. 22, no. 4, April 2017, http://www.firstmonday.org/ojs/index.php/fm/article/view/6872/6089

Winner, L. (1980): 'Do artifacts have politics?', *Daedalus*, vol. 109, pp. 121-136.

Wright, K. B., and Bell, S. B. (2003): 'Health-related support groups on the Internet: Linking empirical findings to social support and computer-mediated communication theory', *Journal of Health Psychology*, vol. 8, no. 1, January 2003, pp. 39-54.

Wolfgang Prinz (2018): Blockchain and CSCW – Shall we care? In: Proceedings of the 16th 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/ecscw2018_13

Blockchain and CSCW – Shall we care?

Wolfgang Prinz
Fraunhofer FIT, Schloss Birlinghoven, 53754 Sankt Augustin
wolfgang.prinz@fit.fraunhofer.de

Abstract. This exploratory paper examines the relationship between CSCW and emerging blockchain technologies. Although the blockchain technology is at first sight not directly related to CSCW, this paper will identify a number of CSCW research areas that are relevant and that can either profit or contribute to blockchain research. To open CSCW research to new areas and to stipulate a discussion between the disciplines, the paper will start with a brief introduction to basic blockchain concepts followed by an exploration of the relationships between the two research areas. It concludes with an initial proposal on how CSCW research results and concepts can inform blockchain design.

Introduction

Over the last two years, we have experienced an increasing interest into blockchain technologies. This is furthermost caused by the hype around cryptocurrencies such as Bitcoin or Ether. However, more interesting than this fascination in financial speculations are the promises of the technology that underlie all these cryptocurrencies, which is the blockchain technology.

In 2008 Satoshi Nakamoto described in his white paper (Satoshi Nakamoto 2008) the basic principles of a blockchain infrastructure. Although the first implementation was up and running already in January 2009 it took a long time, until approximately 2015, before the blockchain and cryptocurrency idea received an uptake by a larger community as well as the general public. Nowadays blockchains are not just considered as a new technology but also as the enabler for a new generation of a WWW or internet of trust, i.e. the 4th generation after the

internet of information, followed by the internet of service and the internet of things (Iansiti and Lakhani 2017).

This assumption is based on the following properties of a blockchain:

- The blockchain technology enables a consensus building within a network of peers instead of consensus provision by an intermediary or a central platform. Thus, it enables a trustful cooperation within a network without the need of a centralized authority that provides a trusted stakeholder service.
- All transactions stored and managed by a blockchain are irreversible and comprehensible. This makes a blockchain very suitable to store transaction data that must be auditable and it can support cooperation processes between network partners who do not trust each other per se.
- A blockchain enables the transfer of values and rights without the need of a trusted 3rd party, thus it provides a notary functionality
- Smart Contracts, which are code snippets that are an integral part of a transaction enable the execution of "contracts" between cooperation partners and may form a basis for a shareconomy.

Interestingly, some of these properties are also relevant for cooperation support applications or touch upon CSCW research. The next section explores these aspects.

Relationship of Blockchain and CSCW Technologies

Proof of Work and Consensus building versus Operation Transformation

A blockchain network is built upon a peer-to-peer network of so-called mining nodes. These nodes exchange information about submitted transactions, but they also compete for the right to validate transactions in a new block that is accepted by the other nodes and then added as a new block to the global blockchain (Christidis and Devetsikiotis 2016). In most of the current blockchain implementations this consensus building is based on so called proof of work algorithms (Abadi et al. 2005). One aim of this algorithm is to ensure that the mining nodes that compute the next block are randomly selected by solving a crypto puzzle. An important property of this crypto puzzle is that it is asynchronous, i.e. it requires a certain amount of work to solve the puzzle, but once it is solved, other partners can quickly check that the solution is correct. Since this approach is not only time consuming but also very energy consuming, a lot of research is undertaken to find alternative methods such as proof of stake or lottery based approaches (Prinz et al. 2018). Nevertheless, all of these approaches aim at the selection of a single mining node that gets the right to serialize all transactions in a

block for being added to the global blockchain. Furthermore, the approach must guarantee that none of the network nodes is able to add manipulated information. In distributed systems this problem is also known as the Byzantine generals problem (Lamport et al. 1982).

CSCW research has dealt with a similar problem in the context of shared editing. First solutions have already been presented in (Ellis and Gibbs 1989) and later in (Ellis et al. 1991) with the development of operation transformation algorithms such as Grove. Until today this research strand is active within the CSCW community (MacFadden et al. 2017). In fact, all these operation transformation approaches can be considered as a consensus building approach and thus they become relevant in a blockchain context.

The main difference between the blockchain and the CSCW approach is the consensus finding approach. Blockchain consensus is based on a competition between the networked nodes. Once a node has identified a possible serialization of transactions, it is checked and accepted by the other network partners. Operation transformation methods achieve consensus about the correct serialization of transactions by a distributed algorithm that takes the context of the transaction origin into account, e.g. by using state vectors.

Therefore, CSCW research may be able to contribute new solutions or even early day solutions such as (Dourish 1996) towards consensus building in a blockchain. This can be based on a proof of collaboration awareness, using a distributed algorithm that validates transactions based on their cooperation context, such as operation transformation. The goal should be to overcome the current limitations inherit to the proof of x algorithms with respect to performance and scalability.

Irreversible Transactions, Smart Contracts and Workflow systems

The blockchain data structure, in combination with consensus building methods, ensures that transactions stored in a blockchain become irreversible. For notary-like applications, this is an essential prerequisite. Since smart contracts are an integral part of a transaction, they become as irreversible as the transaction itself.

Experiences with the development of CSCW systems have already taught us from the very early days that successful cooperation systems can only be build using participatory and user centric design methods (Prinz et al. 1998), (Holtzblatt and Jones 1993).

This raises the dilemma between irreversible software and an evolutionary design approach. One solution might be to develop methods to check the correctness of smart contracts (Osterland, Thomas and Rose, Thomas 2017). However, even a correct smart contract may become problematic if the organizational context or cooperation environment was changed or developed towards a new direction. Therefore, smart contracts must be adoptable to new regulations, environmental conditions or even exceptions.

This discussion is similar to the early discussion with respect to workflow systems and their ability to support exceptions (Kreifelts et al. 1991), to evolve with changing organizations, or to support vage cooperation processes (Herrmann and Loser 1999). As a suitable answer to these problems, CSCW research often followed the path of providing a cooperation media instead of a predefined cooperation process (Bentley and Dourish 1995), (Gräther et al. 1997).

Applying this approach to the design of smart contracts would result in the following guideline: smart contracts should only represent very simple building blocks governed by more flexible coordination systems that can be adopted to organizational requirements. This would result in a two layer based approach. Smart contracts build the bottom layer of irreversible building blocks, while the coordination systems support flexibility by the orchestration of the smart contracts. Smart contracts in this sense can be compared with basic cooperation patterns (Aalst et al. 2003), (Martin and Sommerville 2004), (Herrmann et al. 2003) that can be rearranged flexibly to support more complex cooperation scenarios (Prinz et al. 2009). The Freeflow approach presented in (Dourish et al. 1996) follows a similar direction. This implies that a decentralized autonomous organization (DAO[1]) is no longer being fully build und established by smart contracts but by a combination of smart contracts with a cooperation layer on top. This can add flexibility, but one should also be aware that such a DAO becomes modifiable not just to evolve with an organization but also it can be manipulated for malicious reasons. A solution to overcome this problem is a versioning of this cooperation layer in combination with securing the integrity of the layer by a representation as a transaction in a blockchain. This enables a proof of the invulnerability of the cooperation layer by comparing its current status with its representation in the blockchain.

Blockchain and Reputation Management as a Foundation for a Shareconomy

Beside managing transactions, a blockchain is also suitable to manage identities. The combination of identity management with irreversible smart contracts provides the basis for trusted cooperation processes. This is because, as soon as two cooperation partner have agreed on a smart contract to manage the cooperation transactions, nobody will be able to change this smart contract later on. For example, if the smart contract prescribes a payment after a particular transaction in the blockchain, this payment will be initiated automatically and cannot be obstructed by any of the involved partners. Thus, people can trust in the exact completion of the agreement encoded in the smart contract.

In CSCW and community systems this trust is often achieved by recommendation and reputation management systems (Collier and Hampshire

[1] https://en.wikipedia.org/wiki/Decentralized_autonomous_organization

2010). The rise of platforms such as Airbnb is also an indicator that people are willing to share if someone provides a secure and trustful process. Platforms can achieve this by providing a reputation management and by taking risks (loss of payment, etc.) away from the users. A blockchain replaces this organizational trust management by the algorithmic management of the process. Nevertheless, a blockchain can also support the traditional reputation management approach that is often based on user recommendations such as "likes" by a comprehensible provision of a user's transaction history.

In summary, a combination of CSCW technologies (reputation and recommendation) with the inherent properties of a blockchain can be a solid foundation for decentralized shareconomy networks.

Affordances of a Blockchain

In the early 90's CSCW researchers explored the opportunities of media spaces for the cooperation and awareness support of dislocated users (Mackay 1999). An important aspect that media space research introduced into CSCW is the concept of affordances (Gibson 1986). In (Gaver 1992) the affordances of media spaces are described as: "Media spaces convey visual and auditory information between arbitrary points, and thus afford remote collaboration".

The blockchain technology is considered as the enabler of the internet of trust (Tapscott and Tapscott 2016). The following items are a result of applying the affordance concept ("properties of a technology that offer actions to appropriate organisms") to blockchain properties.

A blockchain

- makes transactions irreversible, thus affords comprehensiveness,
- support consensus in a network, thus affords community based agreement,
- enables the transfer of values and rights without the need of a trusted 3rd party, thus it provides a network based notary functionality,
- supports autonomous actions by smart contracts, thus affords coordination.

In summary, we can argue that a blockchain conveys comprehensiveness, community based agreement, a notary function and coordination and thus affords trust and even more provides the basis for a decentralized autonomous organization.

Summary and Conclusion

This brief exploratory paper identifies in a first approach interesting relationships between blockchain and CSCW technologies. The following table summarizes the

findings and proposes a combined approach (CSCB$_{lockchain}$) that aims at combining the better of the two worlds to overcome problems of flexibility, scalability and adoption.

	Blockchain	**CSCW**	**CSCB**
Consensus building	Proof of work / proof of Stake; competition based approaches by means of local algorithms	Collaborative editing: operation transformation; context based approaches by means of distributed algorithms	Proof of collaboration awareness based on a distributed algorithm that validates transactions based on their cooperation context.
Smart contracts	Irreversible program code as integral part of a transaction	Participative and user centric design requires agile software engineering methods.	Understanding smart contracts as cooperation patterns that are governed by versioned and flexible cooperation rules
Reputation and trust	Irreversible transaction records and smart contracts	User reputation and recommendation	Uptake of the transaction records into reputation management approaches. Building blocks for decentralized shareconomy networks.
Affordance	Affordance for trust and comprehensibility	Affordance for seamless cooperation	Affordance for trusted cooperation in a decentralized network

The aim of this paper is to initiate a discussion between the communities of blockchain and CSCW research. Up to now, the communities are disjoint although they both deal with similar cooperation related topics and research questions. Beyond those identified in this paper, further relationships are of interest such as the connection between cryptocurrency mining and reward or recommendation system, or transaction forensics and social network analysis. This paper shall be a starting point to initiate and stipulate this discussion at ECSCW 2018.

References

Aalst, W.M.P.V.D., Hofstede, A.H.M.T., Kiepuszewski, B., Barros, A.P.: Workflow Patterns. Distrib Parallel Databases. 14, 5–51 (2003)

Abadi, M., Burrows, M., Manasse, M., Wobber, T.: Moderately Hard, Memory-bound Functions. ACM Trans Internet Technol. 5, 299–327 (2005). doi:10.1145/1064340.1064341

Bentley, R., Dourish, P.: Medium versus mechanism: Supporting collaboration through customization. In: Marmolin, H., Sundblad, Y., and Schmidt, K. (eds.) Fourth European Conference on Computer Supported Cooperative Work (ECSCW '95). pp. 133–148. Kluwer, Stockholm, Sweden (1995)

Christidis, K., Devetsikiotis, M.: Blockchains and Smart Contracts for the Internet of Things. IEEE Access. 4, 2292–2303 (2016). doi:10.1109/ACCESS.2016.2566339

Collier, B.C., Hampshire, R.: Sending Mixed Signals: Multilevel Reputation Effects in Peer-to-peer Lending Markets. In: Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work. pp. 197–206. ACM, New York, NY, USA (2010)

Dourish, P.: Consistency Guarantees: Exploiting Application Semantics for Consistency Management in a Collaboration Toolkit. In: Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work. pp. 268–277. ACM, New York, NY, USA (1996)

Dourish, P., Holmes, J., MacLean, A., Marqvardsen, P., Zbyslaw, A.: Freeflow: Mediating Between Representation and Action in Workflow Systems, (1996)

Ellis, C.A., Gibbs, S.J.: Concurrency Control in Groupware Systems. SIGMOD Rec. 18, 399–407 (1989). doi:10.1145/66926.66963

Ellis, C.A., Gibbs, S.J., Rein, G.L.: Groupware: Some issues and experiences. Commun. ACM. 34, 39–58 (1991)

Gaver, W.W.: The Affordance of Media Spaces for Collaboration. In: Turner, J. and Kraut, R. (eds.) CSCW '92: Conference on Computer Supported Cooperative Work - Sharing Perspectives. pp. 17–24. ACM Press, Toronto, Canada (1992)

Gibson, J.J.: The Ecological Approach to Visual Perception. Lawrence Erlbaum Associates, Publishers, Hillsdayle, New Jersey (1986)

Gräther, W., Prinz, W., Kolvenbach, S.: Enhancing Workflows by Web-Technology. In: Hayne, S. and Prinz, W. (eds.) GROUP'97: International ACM SIGGROUP Conference on Supporting Group Work. pp. 271–280. ACM Press, Phoenix, AZ (1997)

Herrmann, T., Hoffmann, M., Jahnke, I., Kienle, A., Kunau, G., Loser, K.-U., Menold, N.: Concepts for Usable Patterns of Groupware Applications. In: Proceedings of the 2003 International ACM SIGGROUP Conference on Supporting Group Work. pp. 349–358. ACM, New York, NY, USA (2003)

Herrmann, T., Loser, K.-U.: Vagueness in models of socio-technical systems. Behav. Inf. Technol. 18, 313–323 (1999)

Holtzblatt, K., Jones, S.: Contextual Inquiry: A Partcipatory Technique for System Design. In: Schuler, D. and Namioka, A. (eds.) Participatory Design: Principles and Practices. pp. 177–210. Lawrence Erlbaum Associates, Hillsdale, NJ (1993)

Iansiti, M., Lakhani, K.R.: The truth about blockchain. Harv. Bus. Rev. 95, 118–127 (2017)

Kreifelts, T., Hinrichs, E., Klein, K.-H., Seuffert, P., Woetzel, G.: Experiences with the DOMINO Office Procedure System. In: Bannon, L., Robinson, M., and Schmidt, K. (eds.) Second European Conference on Computer-Supported Cooperative Work. pp. 117–130. Kluwer Academic Publishers, Amsterdam, Netherlands (1991)

Lamport, L., Shostak, R., Pease, M.: The Byzantine Generals Problem. ACM Trans Program Lang Syst. 4, 382–401 (1982). doi:10.1145/357172.357176

MacFadden, M.S., Ng, A., Ignat, C.-L., Gu, N., Sun, C.: The Fifteenth International Workshop on Collaborative Editing Systems. In: Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. pp. 351–354. ACM, New York, NY, USA (2017)

Mackay, W.E.: Media Spaces: Environments for Informal Multimedia Interaction. In: Beaudouin-Lafon, M. (ed.) Computer Supported Cooperative Work. pp. 55–82. John Wiley & Sons Ltd. (1999)

Martin, D., Sommerville, I.: Patterns of Cooperative Interaction: Linking Ethnomethodology and Design. ACM Trans Comput-Hum Interact. 11, 59–89 (2004). doi:10.1145/972648.972651

Osterland, Thomas, Rose, Thomas: Correctness of Smart Contracts for Consistency Enforcement, (2017)

Prinz, W., Jeners, N., Ruland, R., Villa, M.: Supporting the Change of Collaboration Patterns by Integrated Collaboration Tools. In: Camarinha-Matos, L.M., Paraskakis, I., and Afsarmanesh, H. (eds.) Levaraging Knowledge for Innovation in Collaborative Networks - 10th IFIP Working Conference on Virtual Enterprises, PRO-VE 2009. pp. 651–658. Springer, Thessaloniki, Greece (2009)

Prinz, W., Mark, G., Pankoke-Babatz, U.: Designing Groupware for Congruency in Use. In: Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work. pp. 373–382. ACM, New York, NY, USA (1998)

Prinz, W., Rose, T., Osterland, T., Putschli, C., Osterland, T., Putschli, C.: Blockchain. In: Neugebauer, R. (ed.) Digitalisierung: Schlüsseltechnologien für Wirtschaft und Gesellschaft. pp. 311–319. Springer Berlin Heidelberg, Berlin, Heidelberg (2018)

Satoshi Nakamoto: Bitcoin: A Peer-to-Peer Electronic Cash System, https://bitcoin.org/bitcoin.pdf

Tapscott, D., Tapscott, A.: Blockchain Revolution: How the Technology Behind Bitcoin Is Changing Money, Business, and the World. Brilliance Audio (2016)

Fetter, M.; Müller, A.-L.; Vasilyev, P.; Barth, L. M.; Gross, T.: Towards a Better Understanding of Availability and Interruptibility with Mobile Availability Probes. In Proceedings of 16th European Conference on Computer-Supported Cooperative Work - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2018_14

Towards a Better Understanding of Availability and Interruptibility with *Mobile Availability Probes*

Mirko Fetter, Anna-Lena Müller, Petr Vasilyev, Laura Marie Barth, Tom Gross
Human-Computer Interaction Group, University of Bamberg, Germany
(hci(at)uni-bamberg.de)

Abstract. In cooperative work shared awareness on mutual availability is important for the overall performance of the team. There has been great research on quantitatively analysing users' behaviour and automatically detecting their interruptibility. In this paper we present our approach towards a better qualitative understanding of availability of users. Leveraging on experience sampling and cultural probes we developed a mobile tool to collect Mobile Availability Probes. We motivate the need for a better qualitative understanding of availability, introduce our approach and the Mobile Availability Probes, and present and discuss initially collected availability data.

Introduction

Interruptibility is a vital research topic in computer-supported cooperative work (CSCW) and human-computer interaction (HCI) research. It can be broadly defined as the condition of being willing and able to handle interruptions—even if this interruption might be disturbing an active process. Particularly with the introduction of notifications as a prime interaction mechanism in current smartphones (Sahami Shirazi *et al.* 2014), the topic recently started to attract a wider audience. In the realm of awareness research (Gross 2013) there has been an on-going discourse on how to optimise the balance between the benefits of being informed and the costs of being disrupted (Hudson & Smith 1996). Furthermore, understanding the use of computer-mediated communication (CMC)

technologies and their implications for users' interruption is a central and constant theme.

Especially instant messaging (IM)—which introduced a novel, brief and spontaneous communication style—became a prime research strand (Cutrell *et al.* 2000) for understanding interruptibility. The more holistic perspective that often underlies CSCW research goes beyond a binary distinction of being interruptible or not interruptible. Also, researchers started looking beyond the assessment of the "physiological ability to switch focus" or the "cognitive affect on task performance", and began to investigate the "user sentiment" towards interruptions (Turner *et al.* 2015, p. 802) as well. The notion of *availability* in IM and beyond promotes a more dyadic perspective on interruptibility that aims at considering attitudes towards the communication of the recipient as well as the sender—the interruptee and the interrupter. Managing one's availability in CMC is a complex act (Birnholtz *et al.* 2010) that goes far beyond managing one's general interruptibility, as it includes aspects of social roles and norms, and their individual interpretation and resulting expectations. For instance, it might make a fundamental difference if during a work meeting a user of a mobile phone—and, additionally, other attendees of the meeting—is interrupted by an important urgent message of a family member versus by a notification on an outstanding software update. Thus, when looking at recent research on predicting interruptibility, the majority of current research seems to target at a simplified conceptualisation of interruptibility in order to be able to better quantify and statistically compute interuptibility, yet at the same time neglects these nuances (Turner *et al.* 2015).

When asking users, *availability* is often inverted in the sense that it is explained with examples of personal *un*availability (e.g., talking about the personal unavailability while participating in a meeting, operating a car, or writing a report on a computer). Rarely, examples are given that describe situations in which persons are explicitly available. We believe that such a bias also restricts the potential for analysing the solution space. Moreover, designers and developers of interruption technology often focus on *un*availability when developing sensors that capture data that might be indicators for unavailability (e.g., capture voice activity to infer a person is speaking) or non-interruptibility (e.g., analyse calendar entries to infer a person is in a lecture). It seems that we developed a workable understanding of when people are not available, but do not fully understand what good indicators are for a person being available. Another reason, is that non-interruptibility often seems to be more generalisable and absolute, and therefore more graspable, while availability tends to be more selective and fragmented—for example towards different contacts (Fetter *et al.* 2010). Therefore, we see a research opportunity for developing a better understanding of how to support humans managing their availability. A key challenge thereby is the question, on how to better research interruptibility and availability in daily life. Both concepts

seem to be very volatile, erratic, and unobservable for an external person (Avrahami *et al.* 2007).

A prominent approach has been to use the *Experience Sampling Method* (ESM) (Hektner *et al.* 2006) in order to collect self-reports of users on their current interruptibility or availability (Fetter *et al.* 2010; Fetter *et al.* 2011; Horvitz & Kapoor 2008; Hudson *et al.* 2002; Rosenthal *et al.* 2011). The results of such studies are detailed time-series, spanning a few days or weeks, in which participants assess their personal interruptibility or availability on some form of linear scale. Often additional information is logged (Turner *et al.* 2015) in these studies. Sometimes this happens to reflect about the general nature of influencing factors (Hudson *et al.* 2002), but most of the times the logged data is used to simply compute statistical models that should be able to forecast a user's current interruptibility or availability in a given situation.

That said, an undeniable quality of ESM is, that through its repetitive nature it manages to grasp a good cross-section of people in different states of availability. However, the results from many interruptibility and availability studies are often very prosaic and analytical, and fail to grasp the richness of human social interaction. The results allow us to understand the 'if' and 'when', but seldom the 'why' of people's availability, as such studies often miss to record the underlying texture of human life. Other approaches seem to be much better in capturing these underlying textures of daily live, providing new insights and thus allowing for novel perspectives. For example *cultural probes* (Gaver *et al.* 1999b)—small packages filled with maps, postcards, cameras, booklets, and other material, that can be distributed among members of smaller communities to provoke inspirational responses. While others have successfully used cultural probes to break up stereotypes of domestic research (Gaver *et al.* 1999a), we think they can be used to question our preconceptions about availability. However, the data cultural probes deliver is very fragmentary and incomplete (Boehner *et al.* 2007; Gaver *et al.* 1999b).

In our approach of *Mobile Availability Probes* we aim to combine and complement ESM and cultural probes as a means to better understand how people construct their availability. In the following we provide more background information and take a look at related work before lying out our concept. We report on an early exploratory study and reflect on the collected data. From these reflections, we draw our conclusions on the viability of our approach and provide an outlook.

Background and Related Work

In this section we narrow down the term availability and have a closer look at the methods ESM and cultural probes, and reflect on related work.

From Interruptibility to Availability

In the following we outline where the concepts of interruptibility and availability overlap, and then show up where they differ in respect to CMC.

There is a plenitude of definitions for interruption in the literature (e.g., Iqbal & Horvitz 2007; Jett & George 2003; McFarlane & Latorella 2002; Ritter *et al.* 2014) that often broadly defines the term, for example as "a synchronous interaction which was not initiated by the subject, was unscheduled and resulted in the recipient discontinuing their current activity" (O'Conaill & Frohlich 1995, p. 262).

Yet, in order to be able to relate the term interruptibility to the term availability, it is necessary to further break down this broad concept (cf. Figure 1). To achieve this, we categorise the source and nature of the interruption. Two overarching groups in this respect are whether the interruptions are stimulated externally or internally (Mark *et al.* 2005). Thereby *external* interruptions result "from events in the environment" whereby *internal* interruptions come "from our own thought processes—new ideas that draw attention from the current activity." (Miyata & Norman 1986, p. 268). Hence, internal interruptions are "self-initiated" while for external interruptions it is "a condition in the environment that motivates switching" (Gonzalez & Mark 2004, p. 118). In Figure 1 on the highest level we accordingly differentiate between interruptions that origin from an internal source as the "Self" and those that that origin from external sources as for example "another person, computer, other animate object, [or] inanimate object" (McFarlane & Latorella 2002, p. 19).

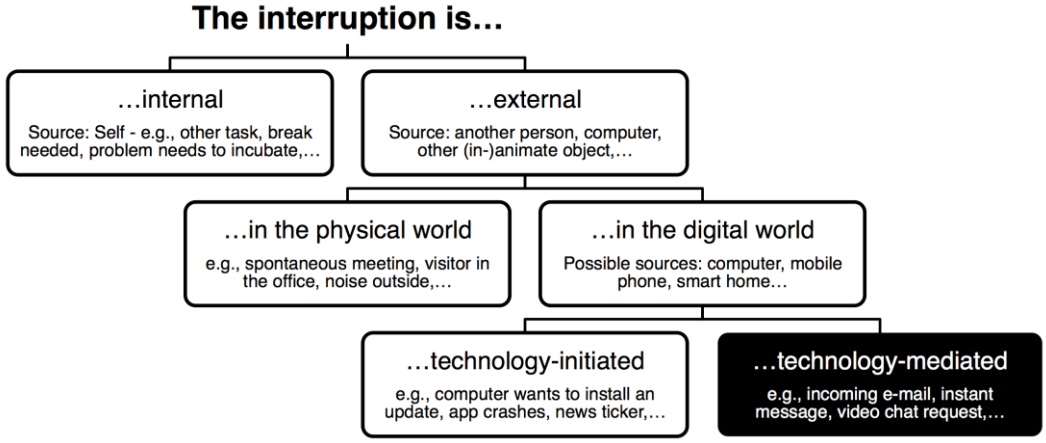

Figure 1. Categorising interruptions based on the source and nature of the interruption.

As this work focuses on *availability*, we further want to break down only external interruptions. We do that by differentiating between interruptions that are originating from the *physical* and those that are originating from the *digital* world. Interruptions in the physical world can have multifarious causes: a colleague

coming into the office with a question, some noise from a construction site outside that makes us close the window, etc. In the digital world, many of the interruptions are originating from notification systems (McCrickard *et al.* 2003b)—that is, hard- and software systems that inform users of events of interest, thus satisfying their multitasking information demands. In this context, an interruption is "an event within the notification system prompting transition of attention focus from a primary task to a notification" (McCrickard *et al.* 2003, pp. 551). Today, such notifications not only originate from computers and phones, but also from in-vehicle information systems in cars, reminding us that we need to go to an inspection, as well as from a smart speaker in our living room, telling us our parcel is out for delivery.

Furthermore, it is important to distinguish between interruptions *initiated* by technology and those *mediated* by technology. The first category often consists of scheduled or automated notifications informing about an outstanding software update, some outstanding maintenance task, a headline from the news app, or an automated newsletter. In many cases the presentation of these notifications is not even time-critical, if it is not a warning or an error. The latter category refers to personal contact mediated through technology such as somebody writing a text message to a mate from the soccer team, sending an email to a customer, or starting a video call with the grandparents. From our perspective, these two categories need to be treated fundamentally differently.

So, when talking about *availability* in CMC, persons wanting to contact each other makes up only a fraction of all interruptions that might occur to users. Yet, this social availability is a very interesting and relevant aspect. It is affected by our social roles, our tasks, our expectations, and the expectations of others. It is often selective towards different audiences, and not towards one singular status (Fetter *et al.* 2010). One of the definitions for availability provided in literature is described as "a state of mind (whether an individual is receptive for communication or not)" (Harr & Wiberg 2008, p. 244). The complexity of availability also becomes evident from a design space analysis of availability sharing systems (Hincapié-Ramos *et al.* 2011) that shows how differently the topic is approached in terms of solutions.

We therefore argue that the topics of interruptibility and availability need to be more disentangled in future research, and availability should not be treated as a specificity of interruptibility.

Experience Sampling Method and Cultural Probes

In the following, the two methods underlying our approach are introduced and discussed. They are very different, but have in common that they capture data *in-situ*—that is, in the moment and not retrospectively.

The Experience Sampling Method (ESM) is a research methodology developed in social psychology (Hektner *et al.* 2006) that has been successfully adapted for

the purpose of research in HCI. Over the course of usually several days or weeks each participant in an ESM study is required to record their inner states, experiences, feelings, or attitudes towards an overarching research question several times a day. Towards this end the participants repeatedly fill out an ESM form—a short questionnaire including anything from open-ended questions to psychometric scales. The method has shown to achieve two things. First, it is able to capture detailed and in-depth data of individual participants through repetition in a form of time-series. Secondly, it is able to capture fine-grained subjective assessments of a person's inner states or feelings in the wild. In HCI and CSCW research it has been applied often to study the interruptibility of different groups of people (Avrahami *et al.* 2007; Horvitz & Kapoor 2008; Hudson *et al.* 2002; Rosenthal *et al.* 2011; Turner *et al.* 2015). While the method generally can be used to collect qualitative data (Hektner *et al.* 2006), the repetition of the same qualitative questions over time often tempts researchers to quantify qualitative data by coding and counting the qualitative answers. Furthermore, such studies can be quite laborious and intrusive (Mehrotra *et al.* 2016) for the participants, which can leads to challenges with drop-outs.

Cultural Probes (Gaver *et al.* 1999b), on the other side, were introduced as a ludic methodology to serve the understanding of certain settings and situations while at the same time embracing the uncertainty and fragmentation of its finding. Study participants receive small probing packages including different materials like postcards, maps, single-use camera, and diaries that aim to "provoke inspirational responses" (Gaver *et al.* 1999b, p. 22) which are later analysed and discussed in interviews, focus groups, or workshops. Originating from a design context, cultural probes were conceived to inspire rather than to inform. Cultural Probes do not aim to find a singular truth, but to provoke novel thoughts and shake up existing preconceptions. In academia and industry this method has been quickly adopted, yet the way it was interpreted often deviates from its intention (Boehner *et al.* 2007; Gaver *et al.* 1999a). A major critique on the adoption is that an originally open and interpretive methodology is often put into a straitjacket of formalism and objectiveness.

Related Work

In many studies in CSCW and HCI interruptibility and availability needs have been analysed. However, there seems to be an underlying trend. Earlier studies often tried to get qualitative insights in order to deeply understand the users' attitudes, needs, and coping strategies. For example, Nardi et al. (2000) qualitatively analysed the IM use of 20 people through interviews and observations, and only supplemented their result with logs of IM messages. Hudson et al. (2002) used an ESM based approach to understand availability and interruptibility attitudes of twelve IBM managers. They also used the results as a

base to conduct qualitative interviews, to develop a deeper understanding of availability.

Today many of the studies use a rather quantitative approach in combination with machine learning (Turner *et al.* 2015). For example, in a large study by Yahoo Japan (Okoshi *et al.* 2017) 680,000 people used an application that detected interruptible moments. However, such approaches are often limited when it comes to capturing the peculiarities of human needs and subtleties of human practises. In many cases all notifications are treated equally—yet, as we already pointed out: a message from a loved one is not treated differently from the request to update a rarely used app.

The basic idea of transferring the concept of mobile cultural probes to mobile phones has been explored before. With Mobile Probes (Hulkko *et al.* 2004) others previously explored the usage of phones to collect qualitative data in a digital manner. The concept of mobile cultural probes was explored in two studies on the two overarching topics of shopping and mobile work. In the shopping study with 13 participants they used a J2ME application to collect the data. In the mobile work study, short messages (i.e., MMS and SMS) were used to send questions to the participants. However, it was only possible for the participants to send text and images. Others have used Digital Cultural Probes (Iversen & Nielsen 2003) in an application that allowed children to collect photos and audio clips on a mobile phone. The material was used to inform the design of digital technology for kids. They concluded that an application is able to motivate kids to spontaneously use it and also commented on the richness of the collected material.

Mobile ESM Probes for Understanding Availability

With our concept of *Mobile Availability Probes* we aim to combine the unremitting persistence of the *ESM* with the ludic and inspiring quality of *cultural probes*—yet not replacing them. Mobile Availability Probes are designed to signal participants at a specific interval to record qualitative data in a format that illustratively captures their current situation and practices, with respect to an overarching research question.

We hope that the combined method is able to record inspirational insights in a specific rhythm, and not only in the few moments a study participant deems something is of particular interest and thus worthwhile reporting. When investigating availability, this is an important quality for two reasons: first, if a person is unavailable, the additional effort of capturing the situation for a study might be too high and therefore participants might skip it; secondly, if a person is available, this situation might not seem relevant from participants' perspective and therefore not reported.

Hence, we moulded our concept of Mobile Availability Probes into an application for studying availability needs in everyday life. The application

notifies the participants throughout their day in a random interval (one random prompt per 90 minutes with min. 15 minutes between two prompts) to answer a short ESM form. Answering the form requires the participants to complete several steps—as depicted in Figure 2—with each step basically representing a single screen.

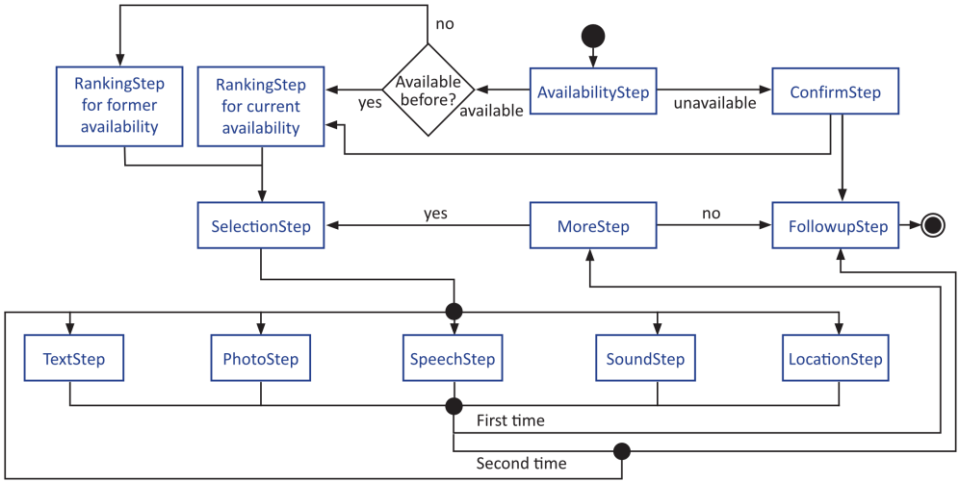

Figure 2. An overview of the participants' path through the ESM form when responding to a sampling request.

First, participants are asked to state their current availability on a scale from "*very available*" to "*very unavailable*" (see *AvailabilityStep* in Figure 3). If participants state that they are *unavailable* or *very unavailable*, we assume answering the full ESM form is inappropriate, yet ask if the person wants to take the survey anyway (*ConfirmStep)*. If the participant decides against taking the survey, they can acknowledge or adapt the time for the next sampling in the *FollowupStep* (cf. Figure 3) and are done with their task. The time for the next sampling is pre-set according to the sampling interval, but can be altered by the participants, if they have a longer period in which they do not want to be interrupted by the ESM Probe.

If the participants in the *ConfirmStep* decide to take the survey despite being unavailable, or while being *very available*, *available*, or *neither available nor unavailable* in the *AvailabilityStep*, they are directed to the *RankingStep* (cf. Figure 3). There, participants are asked to indicate sources of influence on their availability. The question is either related to:

- their current *availability*,
- or if they previously did not fill out a full ESM form, because they were *unavailable* or *very unavailable,* their last *unavailability* (as depicted in cf. Figure 3).

Participants therefore rate the factors that mostly influence—or previously influenced—their availability, on a scale from 0 (no influence) to 9 (very strong

influence). This way, they implicitly rank the factors: *people around them*, their current *location*, their current *task*, or *other factors*.

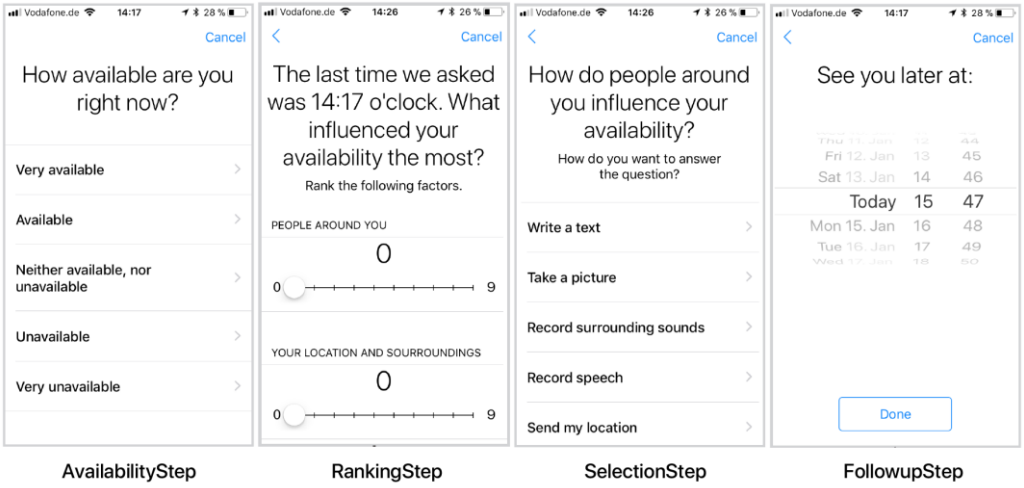

Figure 3. *AvailabilityStep*, *RankingStep*, *SelectionStep,* and *FollowupStep*—screenshots.

After obtaining such a preliminary understanding on how available participants are and what influences their availability, the next steps collect more detailed qualitative insights. In a first step (*SelectionStep*), participants can choose from five different media formats, in which they want to collect the data. They have the possibility to type a short text (*TextStep*), take a photo (*PhotoStep*), record a spoken short text (*SpeechStep*), record surrounding sounds (*SoundStep*), or save a location as GPS coordinates (*LocationStep*). The idea is to allow participants to select the most fitting format in order to capture their current situation with respect to their availability and the influencing factors. The choice of the format is typically determined by individual aspects such as convenience, effort, descriptiveness, social or situational appropriateness, privacy, etc.

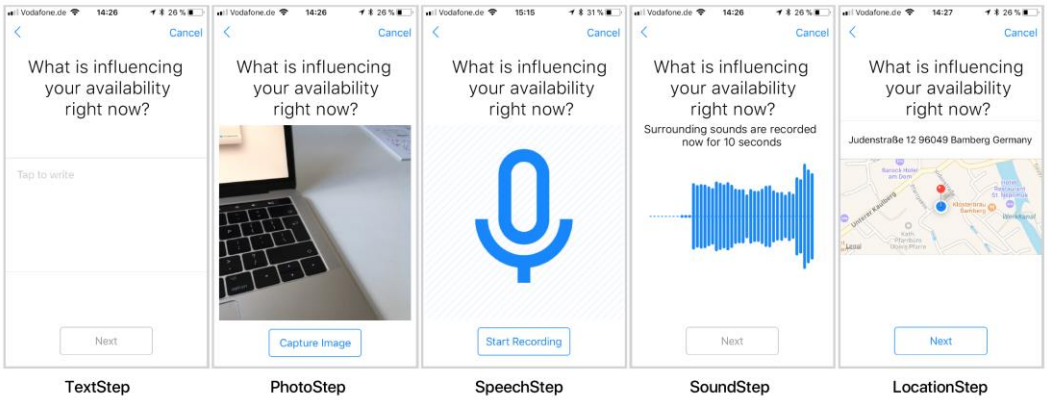

Figure 4. The five possible media steps for recording qualitative answers: *TextStep*, *PhotoStep*, *SpeechStep*, *SoundStep*, and *LocationStep*.

As can be seen in Figure 4, the user interaction with *TextStep* and *PhotoStep* is straight-forward. Participants either type a short text or take a picture with the phone's digital camera. While the *SoundStep* records for a fixed interval of 10 seconds, the recording of the *SpeechStep* is started and stopped by the user. Finally the *LocationStep* uses the current location as default, but allows users to change the recorded location by interacting with the displayed map and pin.

After this step, participants have the possibility to either conclude their ESM form or to choose a second media step in order to collect further qualitative data (cf. Figure 2).

The resulting dataset for each participant comprises snapshots of different moments. The collected meta-data like time, availability, and ranking of the influencing factors help the researcher in analysing and contextualising the qualitative responses in form of written and spoken texts, photos, recorded soundscapes, and locations.

In order to allow the investigator to infer on the collected material, we provide an interactive data exploration tool that allows different views and sorting of the data (e.g., sorting by participant, availability rating, media format). It also provides a detailed look at individual samples (cf. Figure 5 below). It can be used to analyse the data after the study, or to go through the data together with the participant in a post-hoc interview.

Our Mobile Availability Probes concepts was integrated in our application based on the SensQKit—a software framework developed by our group that eases the development of context-aware experience sampling apps based on ResearchKit (Apple Inc. 2018). It was developed for Apple iPhones running iOS 10 or higher. The exploration tool is implemented with Node.js.

Exploratory Study

In order to test the feasibility of the approach, we conducted an exploratory study. Our aim was to investigate, whether our tool and method is able to engage participants in collecting continuously rich data, and to receive some feedback on the tool.

Participants and Procedure

12 participants (6 female and 6 male) between 23 and 56 years old (M = 28.9, SD = 9.0) took part in our study. Nine of them were full-time students, two working in a company, one self-employed. All were recruited through convenience sampling for this pre-study. The study lasted seven days, and the users had at the beginning the possibility to set their personal daily start-time and end-time directly in the application. The times could be chosen without any restrictions. On

an average the participant's choices for times resulted in 11.8 (*SD* =3.3) hours per day, during which they received 42.1 (*SD* = 16.9) sampling requests over the course of the week—so roughly six requests per day in average.

The participants were remotely briefed by reading an instructional PDF document that was sent to them via mail, together with a link to the application. The document also narrowed the conceptualisation of availability, as social availability for all forms of spontaneous computer-mediated communication via smartphones or computers (e.g., instant and mobile messaging, audio and video chats, phone calls).

Seven of the participants used their private iPhones and installed the application via Apple's TestFlight[1]. Five picked up an iPhone 5S with the pre-installed application that we supplied. The participants needed to sign a consent form, clarifying further details on the study and the data usage, directly inside the application before the data collection started. At the end of the study, the users sent the collected data directly from the app via email to us.

Exploratory Results

Overall the participants collected 405 samples. In the following we discuss the general answering behaviour as well as the quality and expressiveness of the collected data.

Answering Behaviour

Participants received 505 sampling notifications in total and reacted to about 86.1% of the notifications. In 29 cases, the participants marked themselves as *unavailable* or *very unavailable* and chose to answer later. This led to an overall number of 405 completed self-reports, ranging from 8 to 66 per user (*M* = 33.75, *SD* = 18.41). Thereby 381 of the reports included one (337) or two (44) qualitative responses (i.e., text, photo, location, etc.). Of the 24 reports that only included the meta-data like the availability assessment but no qualitative responses, 14 were from one participant and the remaining 10 were from 5 participants. Accordingly, half of the participants always used at least one qualitative answer format. 381 of the reports were related to the current availability, while only 24 were related to a previous unavailability. 13 of the 24 answers included a text or a photo, while 11 did not include a qualitative response. The obvious choice for documenting influencing factors of previous unavailability is text. We presume that photos were used, when the general situation did not change much (e.g., watching TV for a longer period of time) and a photo could still be taken at a later point of time.

Table 1. Overview of the use of the qualitative responses.

1 https://developer.apple.com/testflight/

Response Format	Overall	Max. per User	Median per User	Used By
TextStep	276	56	18,5	12 of 12
PhotoStep	69	20	4	11 of 12
SoundStep	40	19	0,5	6 of 12
LocationStep	37	23	0	3 of 12
SpeechStep	3	1	0	3 of 12

Looking at the types of media, we can see that the *TextStep* was used by all participants, and with 276 responses, most often. The *PhotoStep* is the second most used media type; it was used 69 times and by all except one participant. The *SpeechStep* was the least popular format, and only used by three users one-time. While the rationale for including this step was to allow capturing longer descriptive responses that would be too tedious to type in a *TextStep*, we found two of the recorded audio clips were quite similar to the written responses of the *TextStep* (i.e. "I am at home cooking" and "At home, working on some stuff"). The clips had a length of 2 and 3 seconds respectively. The third participant repurposed the *SpeechStep* and treated it like the *SoundStep* to record an 11 second snippet of a lecture. The *SoundStep* with 40 and the *LocationStep* with 37 responses were used almost equally often. While the *SoundStep* was used by half of the participants, only 3 participants used the *LocationStep*.

Quality and Expressiveness of the Material

Analysing the returned data in our exploration tool (cf. Figure 5), showed a wide variety of rich and expressive material. On the one hand, we found material that is confirmative, yet less inspirational. For example, a photo of an unavailable person hurrying to catch a train is more in the line of expected results. On the other hand s, a photo that shows parts of a participant in a bubble bath indicating to be very available stimulates reflections on our conception of what makes up availability. In the same line the texts we received for unavailability more often confirmed our preconceptions (e.g., "Being in a lecture" or "I am working and constantly having customers in front of me[2]"), then those we received for availability (e.g., "Resting after lunch and waiting for the child to finish her nap" or "Tidying up the flat with my roommates").

2 We received written and spoken responses in German as well as in English—for this paper all German responses are translated into English. The translation aims at conserving the content and tonality of the original response.

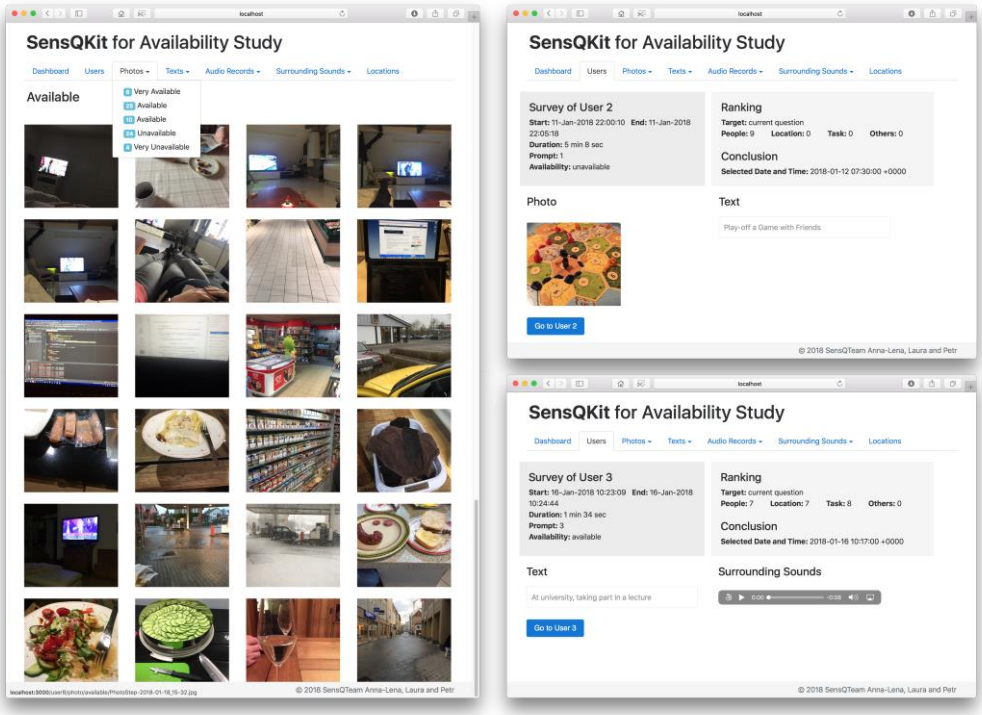

Figure 5. Examples of views of the data exploration tool. Overview with all collected pictures for situations where the participants responded to be available (left); individual sample with meta-data presented together with the collected qualitative data either photo and text (upper right); individual sample with text and surrounding sound (lower right).

We saw reoccurring motives in the photos, texts, and audio responses, like watching television, using a computer, eating or preparing a meal. However, the same motives were associated with rather heterogeneous availability needs—in-between subjects, but also for the same subjects at different times. While many of the received texts were very concrete (e.g., “at work unpacking goods with a colleague”) others were quite vague (e.g., “carnival preparation”) and thus gave much room for interpretation. In the texts we saw the most reoccurring elements, especially when the general situation did not change over several samplings the inserted texts were quite similar.

From the photos we saw that the participants tried to be very privacy preserving—not only with their own privacy, but also that of others. For instance, only one of the 69 photos showed a face of another person; and the body parts (mostly the knees) in the bathroom were also totally anonymous. Yet, by showing the hands, feet, chest area, or backs of co-present people, the participants could still convey the importance of the social interaction for their unavailability or availability. Sometimes the combination of formats (e.g., photos and texts) helped to better understand the reasoning. A close up photo of a board game, with the text “[…] a Game with Friends" hints at the importance of a present person for the current availability. Especially in their photos, it seemed that participants enjoyed

the expressiveness and playfulness of the approach, as for example the depiction of the preparation of a salad showed. Overall the photos tended to provide much richer impressions than the other media types and are also much easier to absorb by the researcher.

The responses from the *SoundStep* were the most difficult to absorb and make sense of. They need to be played one after another, which made them generally harder accessible and it was more difficult to infer the participants' intentions than from other formats. In eight of the recordings it was clear that the person was currently attending a lecture. These recordings perfectly reflected our indented use for the *SoundStep*: to easily capture situations in a socially acceptable manner in which a user might be less interruptible. For other audio clip responses, it was way harder to grasp what is going on—they reached from outside noises to mouse clicking sounds.

Finally, the *LocationStep* was used almost as often as the *SoundStep*, but by fewer users. We assume that the *LocationStep* felt more privacy invading for some of the participants than the *TextStep*, *PhotoStep*, or *SoundStep*. While the other response formats—most prominently photos—also had the potential of being privacy invading, they offered participants more control (e.g., by framing their shots). From the researcher perspective, the pure GPS coordinates—even when displayed on a map—were hard to interpret without further knowledge of the users' significant places and general knowledge of the respective area. Users also revealed locations explicitly. For example, they wrote "[…] at home […]", "[…] at the university […]", "[…] at the gym", "[…] in the office […]", but also various verbal formulations from being on the go. The texts even allowed capturing locations on a more fine-grained level, which would not always be possible with GPS sensor data (e.g., "In the kitchen […]" or "on the sofa"). And also some of the photos revealed details about the current location, such as in a supermarket, the driver's seat in a car, or the passengers' seat in a car.

Conclusion and Future Work

Overall our approach allowed us to collect a considerable amount of expressive and inspirational material. From our first sighting of the material of our exploratory study, it became clear that discussing the captured material with the participants has the potential to convey considerably more information. This can be done either in one-on-one or in focus group sessions. Especially for the locations or recorded sounds, it seems very important to discuss and distil the personal meaning of the recordings with the participants.

The data corroborate our claim that it is far easier to understand what influences *un*availability or even only non-interruptibility than to actually understand when people are available. Yet, with respect to both—availability and *un*availability—the tools helps capturing qualitative responses of potential

meaning to the participants. Going through these responses with participants can shed light on factors that are seen as: (1) one-directional indicators of unavailability (e.g., being in hurry always indicating that one is unavailable); (2) one-directional indicators of availability (e.g., a rest after lunch always indicating that one is available); or (3) bi-directional indicators (e.g., a bubble bath that for the same person on one day is an indicator of unavailability and on the other day is an indicator of availability). In order to build better systems, we are now even more convinced that it is important to understand what determines *availability* above the absence of factors that influence that we are currently not interruptible.

At the moment, the tool and the collected data are primarily used to get a better understanding of the availability of the individual participants as well as availability in general. In the future, it could be extended to be used as a source for training a system that might—after a training phase—better adapt to its user's availability or unavailability.

Acknowledgments

We thank the members of the Cooperative Media Lab, especially Josua Decker, and the participants of the study. Thanks to the reviewers for great comments.

References

Apple Inc. *ResearchKit and CareKit*. https://www.apple.com/researchkit/, 2018. (Accessed 02/02/2018).

Avrahami, D., Fogarty, J. and Hudson, S.E. (2007). Biases in Human Estimation of Interruptibility: Effects and Implications for Practice. In *Proceedings of the Conference on Human Factors in Computing Systems - CHI 2007* (Apr. 28- May 3, San Jose, CA, USA). ACM Press, New York, NY, USA. pp. 51-60.

Birnholtz, J., Guillory, J., Hancock, J. and Bazarova, N. (2010). "on My Way": Deceptive Texting and Interpersonal Awareness Narratives. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work - CSCW 2010* (Feb. 6-10, Savannah, GA, USA). ACM Press, New York, NY, USA. pp. 1-4.

Boehner, K., Vertesi, J., Sengers, P. and Dourish, P. (2007). How HCI Interprets the Probes. In *Proceedings of the Conference on Human Factors in Computing Systems - CHI 2007* (Apr. 28 - May 3, San Jose, CA, USA). ACM Press, New York, NY, USA. pp. 1077-1086.

Cutrell, E.B., Czerwinski, M. and Horvitz, E. (2000). Effects of Instant Messaging Interruptions on Computing Tasks. In *Extended Abstracts of the Conference on Human Factors in Computing Systems - CHI 2000* (Apr. 1-6, The Hague, The Netherlands). ACM Press, New York, NY, USA. pp. 99-100.

Fetter, M., Seifert, J. and Gross, T. (2010). Lightweight Selective Availability in Instant Messaging. In *Extended Abstracts of the Conference on Human Factors in Computing Systems - CHI 2010* (Apr. 10-15, Atlanta, GA, USA). ACM Press, New York, NY, USA. pp. 3817-3822.

Fetter, M., Seifert, J. and Gross, T. (2011). Predicting Selective Availability for Instant Messaging. In *Proceedings of the Thirteenth IFIP TC.13 International Conference on Human-Computer Interaction - INTERACT 2011* (Sept. 5-9, Lisbon, Portugal). Springer, Heidelberg. pp. 503-520.

Gaver, B., Boucher, A., Pennington, S. and Walker, B. (September 1999a). Cultural Probes and the Value of Uncertainty. *interactions* 11, 5. pp. 53-56.

Gaver, B., Dunne, T. and Pacenti, E. (January 1999b). Design: Cultural Probes. *interactions* 6, 1. pp. 21-29.

Gonzalez, V.M. and Mark, G. (2004). "Constant, Constant, Multi-tasking Craziness": Managing Multiple Working Spheres. In *Proceedings of the Conference on Human Factors in Computing Systems - CHI 2004* (Apr. 24-29, Vienna, Austria). ACM Press, New York, NY, USA. pp. 113-120.

Gross, T. (2013). Supporting Effortless Coordination: 25 Years of Awareness Research. *Computer Supported Cooperative Work: The Journal of Collaborative Computing and Work Practices* 22, 4-6. pp. 425-474.

Harr, R. and Wiberg, M. (May 2008). Lost in Translation: Investigating the Ambiguity of Availability Cues in an Online Media Space. *Behaviour & Information Technology* 27, 3. pp. 243-262.

Hektner, J.M., Csikszentmihalyi, M. and Schmidt, J.A. (2006). *Experience Sampling Method: Measuring the Quality of Everyday Life*. Sage Publications, Inc., Thousand Oaks, CA, USA.

Hincapié-Ramos, J.D., Voida, S. and Mark, G. (2011). A Design Space Analysis of Availability-Sharing Systems. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology - UIST 2011* (Oct. 16-19, Santa Barbara, CA, USA). ACM Press, New York, NY, USA. pp. 85-96.

Horvitz, E. and Kapoor, A. (2008). Experience Sampling for Building Predictive User Models: A Comparative Study. In *Proceedings of the Conference on Human Factors in Computing Systems - CHI 2008* (Apr. 5-10, Florence, Italy). ACM Press, New York, NY, USA. pp. 657-666.

Hudson, J.M., Christensen, J., Kellogg, W.A. and Erickson, T. (2002). "I'd Be Overwhelmed, But It's Just One More Thing To Do": Availability and Interruption in Research Management. In *Proceedings of the 2002 SIGCHI Conference on Human Factors in Computing Systems - CHI 2002* (Apr. 20-25, Minneapolis, Minnesota, USA). ACM Press, New York, NY, USA. pp. 97 - 104.

Hudson, S.E. and Smith, I. (1996). Techniques for Addressing Fundamental Privacy and Disruption Tradeoffs in Awareness Support Systems. In *Proceedings of the 1996 ACM Conference on Computer-Supported Cooperative Work - CSCW 1996* (Nov. 16-20, Boston, Massachusetts, USA). ACM Press, New York, NY, USA. pp. 248-257.

Hulkko, S., Mattelmaki, T., Virtanen, K. and Keinonen, T. (2004). Mobile Probes. In *Proceedings of the Third Nordic Conference on Human-Computer Interaction - NordiCHI 2004* (Oct. 23-27, Tampere, Finland). ACM Press, New York, NY, USA. pp. 43-51.

Iqbal, S.T. and Horvitz, E. (2007). Disruption and Recovery of Computing Tasks: Field Study, Analysis, and Directions. In *Proceedings of the Conference on Human Factors in Computing Systems - CHI 2007* (Apr. 28- May 3, San Jose, CA, USA). ACM Press, New York, NY, USA. pp. 677-686.

Iversen, O.S. and Nielsen, C. (2003). Using Digital Cultural Probes in Design with Children. In *Proceedings of the Conference on Interaction Design and Children - IDS 2003* (Jul. 1-3, Preston, England). ACM Press, New York, NY, USA. pp. 154-154.

Jett, Q.R. and George, J.M. (July 2003). Work Interrupted: A Closer Look at the Role of Interruptions in Organizational Life. *Academy of Management Review* 28, 3. pp. 494-507.

Mark, G., Gonzalez, V.M. and Harris, J. (2005). No Task Left Behind? Examining the Nature of Fragmented Work. In *Proceedings of the Conference on Human Factors in Computing Systems - CHI 2005* (Apr. 2-7, Portland, OR, USA). ACM Press, New York, NY, USA. pp. 321-330.

McCrickard, D.S., Catrambone, R., Chewar, C.M. and Stasko, J.T. (May 2003a). Establishing Tradeoffs that Leverage Attention for Utility: Empirically Evaluating Information Display in Notification Systems. *International Journal of Human-Computer Studies* 58, 5. pp. 547-582.

McCrickard, D.S., Chewar, C.M., Somervell, J.P. and Ndiwalana, A. (December 2003b). A Model for Notification Systems Evaluation—Assessing User Goals for Multitasking Activity. *ACM Transactions on Computer-Human Interaction (TOCHI)* 10, 4. pp. 312-338.

McFarlane, D.C. and Latorella, K.A. (March 2002). The Scope and Importance of Human Interruption in Human–Computer InteractionDesign. *Human-Computer Interaction* 17, 1. pp. 1–61.

Mehrotra, A., Pejovic, V., Vermeulen, J., Hendley, R. and Musolesi, M. (2016). My Phone and Me: Understanding People's Receptivity to Mobile Notifications. In *Proceedings of the Conference on Human Factors in Computing Systems - CHI 2016* (May 7-12, San Jose, CA). ACM, N.Y. pp. 1021-1032.

Miyata, Y. and Norman, D.A. (1986). *Psychological Issues in Support of Multiple Activities*. In Norman, D.A. and Draper, S.W., eds. User Centered System Design: New Perspectives on Human-computer Interaction [(1st ed.). Lawrence Erlbaum Associates, Hillsdale, NJ, USA. pp. 265-284.

Nardi, B.A., Whittaker, S. and Bradner, E. (2000). Interaction and Outeraction: Instant Messaging in Action. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work - CSCW 2000* (Dec. 6-10, Philadelphia, PA, USA). ACM Press, New York, NY, USA. pp. 79-88.

O'Conaill, B. and Frohlich, D. (1995). Timespace in the Workplace: Dealing with Interruptions. In *Companion Proceedings of the Conference on Human Factors in Computing Systems - CHI 1995* (May 7-11, Denver, CO, USA). ACM Press, New York, NY, USA. pp. 262-263

Okoshi, T., Tsubouchi, K., Taji, M., Ichikawa, T. and Tokuda, H. (2017). Attention and Engagement-Awareness in the Wild: A Large-Scale Study with Adaptive Notifications. In *IEEE International Conference on Pervasive Computing and Communications - PerCom 2017* (Mar. 13-17, Kailua-Kona, HI, USA). IEEE Computer Society, Los Alamitos, CA, USA. pp. 100-110.

Ritter, F.W., Baxter, G.D. and Churchill, E.F. (2014). *Foundations for Designing User-Centered Systems-What System Designers Need to Know about People*. Springer-Verlag London Limited, London, UK.

Rosenthal, S., Dey, A.K. and Veloso, M. (2011). Using Decision-theoretic Experience Sampling to Build Personalized Mobile Phone Interruption Models. In *Proceedings of the 9th International International Conference on Pervasive Computing - Pervasive 2011* (June 12-15, San Francisco, CA, USA). Springer, Berlin/Heidelberg, Germany. pp. 170-187.

Sahami Shirazi, A., Henze, N., Dingler, T., Pielot, M., Weber, D. and Schmidt, A. (2014). Large-scale Assessment of Mobile Notifications. In *Proceedings of the Conference on Human Factors in Computing Systems - CHI 2014* (Apr. 26-May 1, Toronto, Canada). ACM Press, New York, NY, USA. pp. 3055-3064.

Turner, L.D., Allen, S.M. and Whitaker, R.M. (2015). Interruptibility Prediction for Ubiquitous Systems: Conventions and New Directions from a Growing Field. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp 2015* (Sep. 7-11, Osaka, Japan). ACM Press, New York, NY, USA. pp. 801-812.

Ludwig, T; Kotthaus, C; Stein, M.; Pipek V; Wulf, V. (2018): Revive Old Discussions! Socio-technical Challenges for Small and Medium Enterprises within Industry 4.0. In: Proceedings of 16th European Conference on Computer-Supported Cooperative Work - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2018_15

Revive Old Discussions! Socio-technical Challenges for Small and Medium Enterprises within Industry 4.0

Thomas Ludwig
Christoph Kotthaus
Martin Stein
Volkmar Pipek
Volker Wulf
University of Siegen
{forename.lastname}@uni-siegen.de

Abstract. We may currently perceive an era of massive digitalization within the sector of manufacturing. Summarized as the 'Industry 4.0' vision—as a complex connection between machines, materials, locations, and companies implemented as fully-automatic cyber-physical systems—the way in which manufacturing has been performed will rapidly change, in theory. In practice, however, the outlined configuration of such a vision is not an appropriate option for small and medium-sized enterprises (SMEs). In particular, SMEs and their employees, with their historically-grown experiences and work capacity, secure economic success and need to be put in the spotlight of Industry 4.0 concepts and technologies. Given that the employee is the central success factor within SMEs, the practical adaption of fully-automated and technology-driven concepts raises a variety of socio-technical issues which need to be addressed. Based on an expert workshop with managers and business consultants of SMEs, an interview study with representatives from the German labor union (IG Metall), and the employers' associations, we present current social issues, areas of conflict, and socio-technical challenges SMEs must face. In this exploratory paper, we summarize several research areas that deserve further attention within the next years and which should be considered when conducting studies on SMEs.

The Fourth Industrial Revolution and its Challenges

The first industrial revolution involved the appropriation of water and steam power for mechanization. It was followed by the second industrial revolution, which used electric power for mass production, and by the third industrial revolution, which introduced electronics and information technology (Schwab, 2016). The fourth industrial revolution (Industry 4.0) encompasses a mix of physical hardware and digital software spheres (Schwab, 2016). Industry 4.0 (also referred to as the Industrial Internet of Things or IIoT) is characterized by an increasingly complex connection between machines, materials, locations, and companies in light of advancing information technology. This interwoven connection will undoubtedly have far-reaching effects on manufacturing, the applied production goods, and the internal organization of and external cooperation between companies.

The vision of Industry 4.0 focuses on promoting the company as part of a dynamic, real-time, optimized, and global cross-company and value-added network (Schwab, 2016). Here, smart and connected manufacturing systems – often called cyber-physical productions systems (CPPS) – are considered a technological approach to the challenges of manufacturing within interwoven supply chains and manufacturing locations. They involve closing the gap between data-, technology- and process-driven manufacturing (Broy & Schmidt, 2014; Rajkumar, Lee, Sha, & Stankovic, 2010). The CPPSs include connected, intelligent production plants, which, (a) link the embedded systems of machinery and equipment with Internet-based infrastructures, (b) gather sensor data, and (c) operate actuators. In this way, these CPPSs are *theoretically* able to control the material, goods, and information autonomously.

Although this theoretical vision of Industry 4.0 and CPPS is gaining more and more influence in industry, its *practical* adaption and its configuration *in practice* are still vague and do not consider the individual characteristics of individual companies. For example, they do not consider the company's branch of industry, the company's position in the supply chain, or the size of the company. In particular, the *mittelstand* (mid-tier businesses), as the largest driver of Germany's economy, faces specific challenges concerning Industry 4.0, as this vision primarily focuses on large-scale enterprises such as the large car industry namely (e.g., VW, BMW, or Daimler). Many small and medium-sized enterprises (SME) operate in niche markets and produce small batches or individual pieces in accordance with individual customer requirements. Here, SMEs rely on the important and historically grown employee-related expertise (Brödner, 1986; Wurhofer, Meneweger, Fuchsberger, & Tscheligi, 2018). Employees and their "work capacity" have been ensuring the economic success within SME since several decades. Consequently, the employees are – and according to the SME, will always be – at the core of innovation processes within the mid-sized sector

(Brödner & Latniak, 2004). The expectations and promises of Industry 4.0 or CPPS also face critical perspectives. Past technological waves, such as computer integrated manufacturing (CIM) in the 1980s, predicted in a similar way that human labor would be largely or even completely automated, which however did not occur (Brödner, 2015). Furthermore, the ironies of automation (Bainbridge, 1983) still seem to represent unsolved problems that require the human actor as a central element (Strauch, 2017). In this exploratory paper, we extend the knowledge gained by Wurhofer et al. (2018), which focuses on a micro-level on the role of the worker within smart factories by adding a macro-perspective on the employee, his/her knowledge and the organization. The worker's expertise and the often extremely low degree of automation (and by implication, the high degree of manual manufacturing) make the *adaptation* of the *theoretically* envisioned Industry 4.0 concepts and CPPSs almost impossible. The mainstream vision of Industry 4.0 and fully-automated CPPS will therefore not be an adequate option for the German *mittelstand*.

This exploratory paper takes steps to establish a critical reflection on the current vision of Industry 4.0. We outline the socio-technical areas of conflict SMEs will face regarding Industry 4.0 by relying on the perspective of implementing digitalization within SMEs as an integrated organizational and technological development that includes employees, departments, and the entire value-added chain (Wulf et al.; Brödner 2015). These areas of conflict affect the operational design of work organizations, technology design, and further qualification. We wish to open research areas for CSCW that need to be negotiated between industry, academia, politics, social partners and labor unions and that accordingly deserve further attention within the next years. These areas should be considered when conducting studies within SMEs.

Background information: The German *Mittelstand*

The German economy – especially the industry sector – is characterized by the typical mid-sized structure with mainly small and medium-sized companies. Approximately 95 percent of more than 1,600 plant construction and engineering companies in North Rhine-Westphalia, the most populous German federal state, have fewer than 500 employees. Over two thirds have even fewer than 100 employees (Grothof, 2015).

SMEs are often global leaders within their numerous niche markets (so-called 'hidden champions'). They therefore represent an important section of the successful German economy. The current status of designing CPPS focuses mainly on large companies and implements a techno-centric, top-down perspective on Industry 4.0. CPPSs are applied to a large extent today in highly automated industrial companies based on highly advanced automation technology for production plants (Stich, Deindl, Jordan, Maecker, & Weber, 2015). With

regard to the German mid-sized sector, the question arises how Industry 4.0 concepts and technologies can be applied to meet the requirements of almost every company included in the supply chain. This involves also those kinds of SME that have a high manual degree of manufacturing. At the same time, on how to create new, efficient and economic successful industrial processes and systems.

The main problem is that the stable global economy – which is characterized by continuous product portfolios, clearly defined markets, and stable customer requirements – is a thing of the past (Nyhuis, Fronia, Pachow-Frauenhofer, & Wulf, 2009). Today, SMEs need to be versatile and innovative to persist in and lead the global market. The flexibility and the innovative capability are based on each employee's innovation potential, which introduces their work capacity to the organization structures for the company's economic advantage. This advantage has to be preserved by means of an employee-oriented work and organization structure within the course of standardization and growing automation. With the impact of digitalization and the vision of Industry 4.0, new socio-technical conflict areas will come into play that require an employee-orientated implementation to maintain the potential of the mid-size sector in Germany.

Workshop and Interview Study

To gain insight into the socio-technical challenges and concerns of SMEs regarding the theoretical vision of Industry 4.0, we conducted a workshop with 15 experts, including SME managers and consultants within two districts in North-Rhine Westphalia (the most populous German federal state). The workshop took place in the course of the 'Siegener Mittelstandstagung' (a popular SME conference located in Siegen) with over 250 participants from local industry. Within the workshop, we introduced the concepts of Industry 4.0 and asked the participants to raise the main challenges with regard to this new wave of digitalization by introducing their concerns and writing the key aspects on a flipchart. After this 'collection' phase, the challenges were discussed and fine-specified by all participants. Later on, we clustered the key aspects into six different conflict fields (I-VI). In addition, six interviews were conducted with the most popular German labor union, IG Metall, and the local employers' association (to gain both perspectives on Industry 4.0). In this way, the participants were able to evaluate the concerns and deepen the challenges and risks of the changing work structures in the context of Industry 4.0.

The conflict fields are summarized in the following. We use the term "conflict field" as the main challenges need to be addressed in the future by a participatory social partner-approach – which needs to be discussed among different actors such as employer associations as well as employee organizations – to adapt and implement Industry 4.0 concepts and technologies for the mid-sized sector in the long run.

Conflict field I: Adaptability of CPS and Rollout Strategies

Current Industry 4.0 concepts and technologies, such as cyber-physical production systems, focus on large-scale production with autonomous modifications of the internal and external supply chain. These concepts cannot simply be adapted by SME, as they operate in niche markets and often create individual pieces in accordance with special customer requirements. Though the 'batch size 1' vision of production is an essential idea of Industry 4.0, SMEs do not see themselves able to realize this vision of fully-automated manufacturing in the immediate future. There are three reasons for their doubt: Firstly, SMEs doubt that the investment in CPPS and Industry-4.0 technologies will amortize within an economically acceptable time span (Wischmann, Wangler, & Botthof, 2014). Secondly, the degree of automation in SMEs is on average currently rather low, which results in a high dependency of employees' expertise, which has grown over the years and cannot easily be externalized and transferred into program code (Bracht, Geckler, & Wenzel, 2011). Thirdly, many leaders of SMEs do not have a comprehensive strategy regarding Industry 4.0 to gain an appropriate maturity level (Schröder, 2017). The smaller the company size, the more likely this is to be the case (Sommer, 2015). This is compounded by the fact that employees make demands on technologies that they know from private use (Richter et al., 2017), which often leads to unauthorized use of private IT, also called "Shadow IT" (Steinhueser et al., 2017). At the same time, the employees' working and innovative capabilities have been an essential guarantee for the market position of SMEs for decades. It is therefore feared that knowledge-incorporated CPPSs – which would operate the production processes autonomously – could obstruct employee-driven innovations.

Besides the technical aspects, intra- and inter-organizational and work-scientific connections must be considered and secured for maintaining, reconstructing, and developing innovation-promoting work and competence structures with a social partner-approach. The question of how technological progress can be created without reducing the innovative capability and work capacity of the employees has been asked ever since computers first appeared in the workplace. It is not clear how a heterogeneous mix of an IT landscape, employees, and CPPSs could harmonize. In particular, partial strategies that allow the individualized introduction of CPPSs in SMEs are missing, which consider, integrate, and promote the organizational and employee structure in a socio-technological manner (Scheer, 2012, 2013). The identification of necessary, specific CPPS components proves to be difficult, which could bring the heterogeneous system landscape of SMEs and the individual employee base into accord. Previous IT-driven guiding principles, such as CIM, have already planted doubt in SME regarding the reliable functionality of the technological networking of production as a basic architectural element within a company. CPPS should

instead support the human reflection and adaptation capacity via mechanical precision and speed – in terms of "intelligence enhancement" (Brödner, 2015). Here, the CSCW community has an excellent tradition of analyzing situated work practices that will support addressing the following topics and design regarding Industry 4.0 within the mid-sized sector:

- fitting into existing technological and social infrastructures,
- technological flexibility and expandability with regard to organizational settings, and
- further qualification/job enrichment with Industry 4.0.

Conflict field II: Employee Qualifications

In the course of Industry 4.0, digitalization initiates organizational (r)evolutions, which are a primarily result of the introduction of digital networking systems. The resulting penetration of work processes with digital work tools and virtual cooperation and information instruments will profoundly change the job profile and requirements of employees. Coordination and cooperation within virtual networks requires more and more competences, such as imagination for working contexts and demands, process logics, and specifics of other stakeholders within the value-added chain (supplier, customer, plant manufacturer and operators etc.). The ability to become quickly familiarized with procedures and processes outside one's own subject area is becoming increasingly important. Work activities are expected to become partially virtualized and reorganized in real-time processes that were previously manually and chronically shifted (Geisberger & Broy, 2012). Supporting cooperative work activities beyond several physical boundaries is one of the core research interests within the CSCW community. Work in horizontal department-, company- and cross-company networks implies new qualifications and requires new approaches. These new approaches have to focus on job-related qualification possibilities and "training on the job" (Jacobs & Bu-Rahmah, 2012). These especially include integrative and comprehensive knowledge (process knowledge, operating competences, the ability to understand unknown working procedures, social and analytical competences, and optimization of role interfaces) that is not subject to the employees' specific activity (Gaiziunas, 2009).

Interdisciplinary work contexts are becoming important, as are the necessary competences. This is bound to the fact that the vertical and horizontal networking of companies and the close informational link between suppliers and customers are increasingly abrogating the common differentiation of manufacturing, service, and administration work. This networking of value-added chains facilitates the development of *hybrid* products that consist of tangible assets and complementary services linked by employees' expertise. Here, traditional work and tasks will

become less divided into specific tasks; they will instead be “hybrid,” like the anticipated company networks. Manual and analog activities tend to decrease in favor of immaterial warranty works.

Consequently, important questions and challenges arise for education policy, intra- and inter-company development, and qualification policy. The specific operational evaluation of how digitalization processes will affect job descriptions is still pending and must be realized operationally. Synchronously developing and implementing technology in accord with the development of competences and the adjustment of qualification policies on these different levels creates an opportunity for the cadence of humans and machines and might enable employees to function as helmsmen of the systems. New initiatives and approaches are therefore necessary (e.g., a certified advanced training on the job).

An ever-increasing technological innovative requires greater efforts by the employees in the form of advanced training. Lifelong learning is becoming more and more important. It requires companies to provide corresponding offers and allows for educational breaks. Job-related qualification possibilities gain importance in connection with Industry 4.0. The challenge lies in the integration of learning and qualification possibilities in working processes; these possibilities should be conceived and realized as part of good working conditions and successful occupational trajectories. Competence development requires analyzing the requirements of a specific task and putting employees in a position to meet them. Joint practices or joint communities play a significant role in the acquisition of occupational competences. Thus, in the context of Industry 4.0, it is necessary to support processes of competence development with network technology. This is also necessary for the purpose of cooperation regarding the preservation and provision of experience and its exchange amongst employees within the company or even beyond. When it comes to facilitating the cooperative appropriation of new technologies, it is therefore important to provide learning opportunities that are at or close to the workplace and are process-integrated ("contextualization"). Such provision increases the work capacity by using and passing on experience-based knowledge and implicit knowledge. The increased importance of experience-based knowledge can subsequently become anchored in tariff structures and the assessments of the staffs’ competences. Within the conflict field, ‘qualification requirements and employee qualification, the following topics need to be discussed in the long term by CSCW research in the field of Industry 4.0 and the mid-sized sector:

- integration of knowledge management and training-on-the-job,
- establishment and support of intra- and inter-company learning communities,
- expansion of intra- and inter-company qualification possibilities via relevant offers (e.g., possibility for part-time education, broadly enabling the use of intra-company qualification), and

- processes to control and adapt needs for further training and respective measures and the plannability of learning careers.

Conflict field III: Human-machine-cooperation

While employee qualification is primarily focused on the expansion of employees' working capacity, complex cyber-physical production systems must, from a technical point of view, offer possibilities for the highly efficient production of goods. Against the backdrop of the pressure implicitly built up by the customer by the increasing diversity of variants, by smaller batch sizes, and by increased product complexity, production planning and control of such partially automated complex processes is becoming increasingly confusing. A variety of environmental parameters of the heterogeneous, integrated production resources is in constant interchange with the production's later characteristics and condition. This applies especially to the modern production such as transforming, joining or cutting, since the mechanical function of a plant and different media and tools are joined during the production process – which leads to a great number of plant conditions and dependent process parameters.

Despite the more flexible production and process design, the partly automated, complex production systems pose a great challenge to their introduction, availability and technological controllability (Brödner, Hamburg, & Kirli, 1997; Munir, Stankovic, Liang, & Lin, 2013). The presentation of relevant influencing factors has to be highlighted in real-time based, complex production processes in internal operational and supra-operational contexts. Complex production procedures and processes need to be analyzed in a timely fashion and provided *in situ* for the employees: e.g., the plant operators or (internal/external) decision-makers. The challenge lies, therefore, in the worker-orientated design of new worker-machine interaction types and in enabling the staff to work within highly connected working environments while still remaining in control. This issue is especially problematic with regard to disruptions and errors within the highly complex production processes (Pipek und Wulf 2009). The challenge is to impart competences to employees to address problems in a specific situational context and thereby adjust and restore a regulated working process – particularly in respect to rather fully automated systems. Current machines do not offer functionalities for systemically evaluating internal and external incidents across systems. They also lack an information system which can appropriately support employees and facilitate a suitable procedure for targeted, fast, and efficient production in such a situation. With research discourses around appropriation (Dourish, 2003), appropriation infrastructures (Draxler & Stevens, 2011) or sociale technologies (Ludwig, Boden, & Pipek, 2017), CSCW has already developed concepts for supporting imparting competences to employees. These concepts have to be transferred to the new wave of digitalization.

German (and many European) companies are committed to controlling, recording, and mastering their production processes themselves (in Germany e.g., the DIN standard DIN EN ISO 9000ff). However, the complexity of the current production systems and the current piercing of CPPS, the fast technological advancement, and the close connection of hardware and software in the field of production – these factors pose an operational challenge to employees who use such complex production technologies (Ludwig, Stickel, Boden, & Pipek, 2014). New user interfaces and support tools are, therefore needed, from a technological point of view, to allow the users to keep up with development, to understand the machines independently, and to use them effectively and efficiently for their own work. These interfaces work with the objective of empowering employees in their work and thereby optimizing the economic efficiency of production processes. Creating such interfaces and the employee qualifications involved in them determines whether small and medium-sized enterprises can establish themselves, or rather survive, in Industry-4.0 orientated value-added chains – especially given a lack of investment resources.

SMEs are afraid that, in the future, 'Industry-4.0 certifications' (or rather 'CPPS-ready-certifications') for partly automated, horizontal, value-added chains will determine whether large enterprises will cooperate with them. Such certificates will, on the one hand, ensure certain quality standards and interoperability; on the other hand, they will cause the problem that SME which do not gain such certificates will be pushed out of the value-added chain. It is therefore necessary to draft standardization and certification processes with the participation of the mid-sized industry sector. In accordance with this, bridges and interface technologies need to be created that allow SMEs to gradually and sustainably add concepts to the scope of Industry 4.0 and in this way to simultaneously facilitate an integrated organizational and technological development (Wulf & Rohde, 1995). Within the conflict field of 'human-machine cooperation', the following topics need to be discussed in the long term by CSCW research in the field of Industry 4.0 and the mid-sized sector:

- improvement of control and operability of complex production plants,
- implementation of cooperative decision structures under real-time conditions,
- new hardware-oriented concepts and socio-technical infrastructures for adapting new technologies, and
- cooperation and standardization questions within inter-connected value-added chain and impulses from social partners in practice.

Conflict field IV: Health Protection and new Flexibility Compromises

The real-time networks of work processes and the use of increasingly more efficient, web-enabled devices, changes the potential for timely access to the

employees and the character of work within the core time of an operation. The close cooperation of organizations and departments in other countries, time zones and sectors also expands the socio-spatial relations. Here, spatial and temporal flexibility is increasingly required from the employees so as to harmonize private and occupational wishes. Scientific studies (Collatz & Gudat, 2011) show that constant availability, regular extra hours, and the tendency to blur private and work-related activities have negative consequences on the physical health and long-term performance of employees. The spread of stress symptoms and physical diseases (e.g., burn out) can also be seen in this context. The digitalization of the economy and the entire world requires a new sensitivity to dealing with staff and demands. It also requires a new joint regulation of working hours – for instance, by means of operating agreements. Limits can be set by means of innovative circumstances for work-life boundaries. In this way, employees avoid improper stress – even in working contexts that are mostly automatic and self-organized.

The challenges require better work conditions and health protection, which harmonizes with the better work-life-balance and more flexible requirements of the companies. In the future, CSCW researchers who examine working conditions and intervene with IT artifacts need to work together with social partners to fulfil the reuqirements of good working conditions. New challenges, new learning behaviours, and new types of interactions will encompass flexible operations at work. The impacts of time-critical activities in working environments are challenged by the increasing provision of real-time information and by the generally complex requirements regarding information processing (multi-tasking, frequent work interruptions, and changing demands, etc.) and the health and performance of the employees. These factors must be evaluated during the introduction of new sensor technology. The objective is to maintain and promote cognitive performance and mental health so as to avoid tendencies toward performance compression. Within the conflict field, ‘health protection and new flexibility compromises’, the following topics need to be discussed in the long term by CSCW research in the field of Industry 4.0 within the mid-sized sector:

- evaluation of boundary blurring potentials (work-life-balance),
- avoiding improper stress,
- adjusted work and health protection, and
- development and implementation of adequate work organization concepts.

Conflict field V: Safety of company data and processes

Within the value-added chains, German SMEs are mostly located in the supplier's position. Industry 4.0 postulates a transparent value-added chain, real-time production tracking, and an interface for external views of the production and thus of one’s own company. This conflict field arises from the fact that such

concepts might, on the one hand, strengthen a supplier's position within the value-added chain, but on the other hand, can also be arbitrarily exchangeable with horizontal value-added networks. For example, German SMEs fear that, in the age of globalization, transparent value-added chains could contribute to their own substitution by foreign manufacturers. In addition, SMEs also fear that transparent processes will reveal the profiles of highly qualified employees - the essential guarantors of company knowledge - and that they could be "headhunted" by cooperation partners through various incentives. Regarding the networked technology itself, they also worry about choosing the wrong standard, as there are currently many different approaches (Schröder, 2017).

When it comes to the level of generated industrial data, opportunities and risks go hand in hand. Industry 4.0 technologies generate sensitive industry data in large amounts. This includes sensor data, product information, delivery details, alarm data, error reports, and test results. On the one hand, sensor data from in-house production units, which is particularly important for manufacturing companies, facilitates significant increases in efficiency, the avoidance of disruption-related downtimes, and innovative services such as the worldwide remote maintenance of machines. On the other hand, German SMEs fear that corporate processes will become transparent due to a lack of data security when using industry 4.0 technologies (Schröder, 2017), and that this will increase the pressure from major customers, which could, among other things, be reflected in prices. Moreover, company secrets may also be lost to competitors, as industrial data discloses sensitive information about what is happening in factories – which includes knowledge about production quantities, production control, and error rates.

Medium-sized companies, in particular, are therefore faced with the question of who owns process data and what protections they enjoy in the event of external access. Is data measured by sensors owned by the machine manufacturer, by the manufacturing company, or by the customer who ultimately pays for the production process? The German legal system currently does not recognize any original data-protection rights. Only physical data storages are directly protected against damage and alteration under criminal and civil law, and natural persons are protected against the illegal handling of their personal data. In addition, there is the protection of company secrets under the unfair competition law (§ 17 UWG), the database-manufacturer's ancillary copyright law (§§ 87a ff. UrhG), and relevant intellectual property rights that protect data in the form of personal intellectual creations (copyright law) or inventions (patent law). However, these protection options are only of limited use for Industry 4.0 applications, as their teleological orientation does not take the data in question (especially machine data) into account and therefore only occasionally captures it. To remain competitive in the face of global competition, it may prove necessary in the medium term to introduce an industrial property right on a German level (if not a

European level), which would allow for a risk-minimized handling of industrial data. The European Commission, for example, is striving to achieve this in the course of its ‘Strategy for a Digital Internal Market for Europe’ (COM 192 PUBLIC 2015), but it should be promoted especially by small and medium-sized enterprises as the largest industrial driver. Along with these legal questions regarding data security, CPPSs also raise general IT security threats for companies, such as the failure of critical IT infrastructures due to failure or sabotage.

As a result, SMEs in particular are currently in need of concrete practical advice regarding data and systems protection, secrecy/know-how protection, and proprietary exploitation rights. The following topics arise within the conflict area of security of company data and processes. They must be dealt with by CSCW research in the long term through Industry 4.0 in SMEs:

- equivalent consideration of hard (e.g., encryption) and soft (e.g., data transparency, legal structures) data-protection aspects,
- visibility analysis of operational knowledge carriers (headhunting risks),
- agreement on data ownership (SME, costumer, or machine manufacturers), and
- further development of the legal protection of data as intellectual property (in-company and politically).

Conflict field VI: Employee data protection

The digitalization of the course of Industry 4.0 poses a completely new challenge for both the IT security of a SME and for employee data protection. Networking via mobile devices and the omnipresence of computers both changes work activities and facilitates the recording of employee behavior and its evaluation by algorithms. The involvement of employees in a continuous flow of information between departments and actors in value-added chains and new mutual information, consultation and negotiation approaches through digital technologies presuppose the collection, storage, evaluation and allocation of technology and employee data. This creates requirements for new approaches to employee data protection. The link between technology and personal data, and the evaluation via algorithms, allows for comprehensive performance and behavior profiles. Such a use of Industry 4.0 concepts is met with skepticism or even rejection by most employees (cf., Hornung and Steidle 2005). Acceptance of Industry 4.0 is possible only by evaluating and taking into account the limits of employee trust. Therefore, employee acceptance thresholds must be taken into account when designing Industry 4.0 by involving employees and employee organizations. A crucial factor here is a sensitive approach to the linking of technological and personal data which complies with the principles of necessity, data economy and earmarking, and which focuses on the participation rights and design proposals of

works councils and employees.The following topics need to be addressed in the long-term by CSCW research for Industry 4.0 and the mid-sized sector within the conflict field of ‘employee data protection’:

- determination and consideration of trust limits by participatory introduction of new technological systems;
- use of personal data in accordance with the principles of necessity, data economy and earmarking; and
- anonymization and pseudonymization of personal data.

Conclusion: People in the Center

Within small and medium-sized companies, employees and their experience and work capacity ensure economic success. Thus, the practical embodiment of fully-automated and technology-driven concepts raises a variety of social-technical and organizational issues for employees. We summarize these issues within six areas of conflict: (1) the socio-technological adaptability of CPPS-oriented concepts, (2) the qualification of employees, (3) human-machine cooperation, (4) occupational safety and health, and (5) the security of a company’s and (6) the protection of an employee’s information. The concrete design of Industry 4.0 for the mid-sized sector has to aim at the employee as the most important work resource and as a joint configuration between the social partners and the employers’ associations. With this view, CSCW must build a research agenda that enhance existing concepts and envisions a cooperative network of humans and machines when focusing on SMEs. To achieve this research agenda, companies, employer organizations, and unions and researchers need to work collaboratively with the objective to integrate organizational and technological development in a manner which includes the employees, the departments, and the entire value-added chain as an answer to mainstream Industry 4.0 concepts and technologies.

With this paper, we do not want to give concrete suggestions or guidelines about how to implement the theoretical vision of Industry 4.0. Instead, we wish to sensitize researchers to current and future areas of conflict, and to ‘arenas’ in which small and medium-sized enterprises, social partners, employer associations, and researchers must negotiate their interests and strategies. We want to encounter pure technological concepts of fully-automated CPPS to have a sophisticated perspective for the later design of Industry 4.0 in practice.

The embodiment of the Industry 4.0 vision requires more practice-oriented studies to develop holistic perspectives regarding SMEs in the light of future intelligent and connected supply chains. New options for manufacturing must also be considered. As the CSCW research has an excellent tradition of analyzing situated work practices, the existing concepts within CSCW are predestined to address the special needs, potentials, concerns, and risks of SMEs and their employees in practice, and it allows for the further development of sustainable Industry 4.0 strategies.

References

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*(6), 775–779. https://doi.org/10.1016/0005-1098(83)90046-8

Bracht, U., Geckler, D., & Wenzel, S. (2011). *Digitale Fabrik: Methoden und Praxisbeispiel*. (VDI-Buch, Ed.). Berlin, Heideberg: Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-540-88973-1

Brödner, P. (1986). Skill based manufacturing vs. "unmanned factory"-which is superior? *International Journal of Industrial Ergonomics*. https://doi.org/10.1016/0169-8141(86)90018-1

Brödner, P. (2015). Industrie 4.0 und Big Data – wirklich ein neuer Technologieschub? In H. Hirsch-Kreinsen, P. Ittermann, & J. Niehaus (Eds.), *Digitalisierung industrieller Arbeit* (1st ed., pp. 231–250). Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG. https://doi.org/http://dx.doi.org/10.5771/9783845263205-232

Brödner, P., Hamburg, I., & Kirli, P. (1997). *Leitlinien zur Einführung neuer technischer Verfahren und Systeme im Betrieb*. *Institut Arbeit und Technik: Jahrbuch 1996/97*. Gelsenkirchen.

Brödner, P., & Latniak, E. (2004). Recent findings on organisational changes in German capital goods producing industry. *Journal of Manufacturing Technology Management*. https://doi.org/10.1108/17410380410535062

Broy, M., & Schmidt, A. (2014). Challenges in engineering cyber-physical systems. *Computer*. https://doi.org/10.1109/MC.2014.30

Collatz, A., & Gudat, K. (2011). *Work-Life-Balance*. Göttingen et al.: Hogrefe Verlag GmbH & Co. KG.

Dourish, P. (2003). The Appropriation of Interactive Technologies: Some Lessons from Placeless Documents. *Journal Computer Supported Cooperative Work*, *12*(4), 465–490. https://doi.org/10.1023/A:1026149119426

Draxler, S., & Stevens, G. (2011). Supporting the collaborative appropriation of an open software ecosystem. *Computer Supported Cooperative Work*, *20*(4–5), 403–448. https://doi.org/10.1007/s10606-011-9148-9

Gaiziunas, N. (2009). *Qualifizierung im Supply-Chain-Management. Vom Einkäufer zum Supply-Chain-Manager*.

Geisberger, E., & Broy, M. (2012). *agenda CPS - Integrierte Forschungsagenda Cyber-Physical Systems*. *acatech STUDIE*. https://doi.org/10.1007/978-3-642-29099-2

Grothof, I. (2015, April 10). NRW-Maschinenbau in Pole Position. *Frankfurter Allgemeine Zeitung*.

Hornung, G., & Steidle, R. (2005). Datenschutz: Biometrie am Arbeitsplatz – sichere Kontrollverfahren versus ausuferndes Kontrollpotential. *Arbeit Und Recht*, *53*(6), 201–207.

Jacobs, R. L., & Bu-Rahmah, M. J. (2012). Developing employee expertise through structured on-the-job training (S-OJT): an introduction to this training approach and the KNPC experience. *Industrial and Commercial Training*, *44*(2), 75–84. https://doi.org/10.1108/00197851211202902

Ludwig, T., Boden, A., & Pipek, V. (2017). 3D Printers as Sociable Technologies: Taking Appropriation Infrastructures to the Internet of Things. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *24*(2). https://doi.org/10.1145/3007205

Ludwig, T., Stickel, O., Boden, A., & Pipek, V. (2014). Towards Sociable Technologies: An Empirical Study on Designing Appropriation Infrastructures for 3D Printing. In *Proceedings of DIS14 Designing*

Interactive Systems (pp. 835–844). Vancouver, Canada. https://doi.org/10.1145/2598510.2598528

Munir, S., Stankovic, J. a., Liang, C.-J. M., & Lin, S. (2013). Cyber Physical System Challenges for Human-in-the-Loop Control. In *The 8th International Workshop on Feedback Computing*. San Jose, California, USA.

Nyhuis, P., Fronia, P., Pachow-Frauenhofer, J., & Wulf, S. (2009). Wandlungsfähige Produktionssysteme - Ergebnisse der BMBF-Vorstudie „Wandlungsfähige Produktionssysteme“. *Wt Werkstattstechnik Online*, *99*, 205–210.

Pipek, V., & Wulf, V. (2009). Infrastructuring: Towards an Integrated Perspective on the Design and Use of Information Technology. *Journal of the Association for Information Systems*, *10*(5), 447–473.

Rajkumar, R., Lee, I. L. I., Sha, L. S. L., & Stankovic, J. (2010). Cyber-physical systems: The next computing revolution. In *47th ACM/IEEE Design Automation Conference (DAC)* (pp. 731–736). https://doi.org/10.1145/1837274.1837461

Richter, A., Vodanovich, S., Steinhüser, M., & Hannola, L. (2017). IT on the Shop Floor -Challenges of the Digitalization of Manufacturing Companies. *BLED 2017 Proc.*, 1–16. https://doi.org/10.18690/978-961-286-043-1.34

Scheer, A.-W. (2012). Die Industrierevolution 4.0 verändert Organisationen und Prozesse! Retrieved October 4, 2015, from http://www.august-wilhelm-scheer.com/2012/08/23/die-industrierevolution-4-0-verandert-organisationen-und-prozesse/

Scheer, A.-W. (2013). *Industrie 4.0*. Satzweiss.com.

Schröder, C. (2017). The Challenges of Industry 4.0 for Small and Medium-sized Enterprises. *Friedrich-Ebert-Stiftung*, 28.

Schwab, K. (2016). The Fourth Industrial Revolution: what it means and how to respond. *World Economic Forum*. https://doi.org/10.1038/nnano.2015.286

Sommer, L. (2015). Industrial Revolution - Industry 4.0: Are German Manufacturing SMEs the First Victims of this Revolution? *Journal of Industrial Engineering and Management*, *8*(5), 1512–1532.

Steinhueser, M., Waizenegger, L., Vodanovich, S., & Richter, A. (2017). Knowledge Management without Management – Shadow IT in Knowledge-intensive Manufacturing Practices. In *Proceedings of the European Conference on Information Systems* (pp. 1647–1662). Guimarães, Portugal.

Stich, V., Deindl, M., Jordan, F., Maecker, L., & Weber, F. (2015). Studie - Cyber Physical Systems in der Produktionspraxis. Wuppertal: Clustermanagement IKT.NRW.

Strauch, B. (2017). Ironies of Automation: Still Unresolved After All These Years. *IEEE Transactions on Human-Machine Systems*. https://doi.org/10.1109/THMS.2017.2732506

Wischmann, S., Wangler, L., & Botthof, A. (2014). *Industrie 4.0 - Volks- und betriebswirtschaftliche Faktoren für den Standort Deutschland (AUTONOMIK INDUSTRIE 4.0)*. (L. K. GmbH, Ed.). Berlin: BMWi.

Wulf, V., & Rohde, M. (1995). Towards an integrated organization and technology development. In *Proceedings of the 1st conference on Designing interactive systems: processes, practices, methods, & techniques* (pp. 55–64). https://doi.org/10.1145/225434.225441

Wurhofer, D., Meneweger, T., Fuchsberger, V., & Tscheligi, M. (2018). Reflections on Operators’ and Maintenance Engineers’ Experiences of Smart Factories. In *ACM International Conference on Supporting Group Work (GROUP)* (pp. 284–296). Sanibel Island, Florida, USA. https://doi.org/10.1145/3148330.3148349

Nolte, A.; Jahnke, I.; Chounta, I.-A.; Herrmann, T. (2018): Supporting Collaboration in Small Volunteer Groups with Socio-Technical Guidelines. In: Proceedings of 16th European Conference on Computer-Supported Cooperative Work - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2018_25

Supporting Collaboration in Small Volunteer Groups with Socio-Technical Guidelines

Alexander Nolte[1,2], Isa Jahnke[3], Irene-Angelica Chounta[1], Thomas Herrmann[4]
[1] University of Tartu, [2] University of Pittsburgh, [3] University of Missouri, [4] Ruhr-University Bochum
Corresponding author: alexander.nolte@ut.ee

Abstract. In this paper, we present a study on group work in which student volunteers from different disciplines worked together to create an augmented reality expedition. The goal of the project was to develop an augmented campus tour for students. The project was successful in delivering the app but through post project interviews we found that volunteers were not satisfied with the process and expressed negative insights. In order to understand this phenomenon, we developed and applied a set of categories for detecting underlying problems in socio-technical processes of volunteer group work. Applying those categories to the aforementioned project allowed us to assess their feasibility. This led to refined categories that can potentially support other volunteer groups to create a suitable socio-technical environment.

Introduction

Volunteer groups provide a large variety of valuable contributions to local communities and society at large. They support elections (Boulus-Rødje and Bjorn, 2015) and sports events (Cuskelly et al., 2006), contribute to larger non-profit organizations (Garner and Garner, 2011), devote their time to open source software (Crowston, 2011) or support online production communities such as Wikipedia (Farzan et al., 2012). Most research investigating volunteer collaboration focuses

on large non-profit organizations or open source and online production communities (Boulus-Rødje and Bjorn, 2015; Crowston et al., 2007). Few studies so far have focused on small volunteer groups that collaborate on dedicated activities such as discussion meetings, food giveaways, social events or creating a product. Small volunteer groups face unique challenges in particular related to the way they coordinate their activities. Larger organizations are typically run by a core group of experienced volunteers that split larger efforts into manageable tasks (Cataldo and Herbsleb, 2008; Liao et al., 2016) or even employ a coordinator to take over those duties (Farrell et al., 1998; Harrison, 1995). Small volunteer groups, however, need to coordinate their activities themselves while working towards their main goal at the same time (Nolte, 2018).

In this paper we present the results of a study on a group of volunteers who collaboratively developed an augmented reality (AR) application for university tours over 11 months. During this time, the volunteers developed, tested and delivered an app prototype for a client who used it along with 130 foreign students. Although the project was successful in developing the app, the volunteers afterwards expressed their frustration about the way they collaborated. They perceived it to be chaotic and unorganized, they were not satisfied with the process and they demonstrated negative associations.

This prompted us to look deeper into the process from a socio-technical perspective. Our goal was to identify potential sources of those negative associations as well as reasons for why they still decided to finish the project together despite those negative associations. We thus aim to answer the following questions:

- Why did the volunteers continue the project?
- What were the problems as perceived by the volunteers?
- How can those problems be detected more efficiently?

The last question is of particular importance since volunteers are commonly motivated by a specific cause to which they aim to contribute (Clary et al., 1992; Cobb et al., 2014; Karr and Meijs, 2006). Activities related to analyzing the way they collaborate might thus be eventually perceived as additional work leading to frustration and eventual drop-out. To identify suitable means to detect problems efficiently we developed a set of analytic categories based on existing literature. Then, we used those categories to analyze the results of an interview study which was conducted after the project. The study focused on technical means of support and communication media, individual and group goals and the collaborative process as a socio-technical setting. Based on the analysis we refined those categories so that they can serve as a guide for volunteer groups to detect problems in the way they coordinate. We envision this to enable volunteer groups to improve collaboration

Developing Categories to Evaluate Socio-Technical Collaboration

The aim of this paper is to develop and apply a set of categories that can support small groups of volunteers to detect problems in the way they collaborate. The field of socio-technical systems seems to be a natural starting point for this since the groups we analyze need to coordinate their activities, which involves the use of technology to communicate and share artifacts. Researchers in this field have developed approaches that support collaboration through technological and organizational change since its inception in the early 1950s (Cherns, 1987; Clegg, 2000; Eason, 2005; Fischer and Herrmann, 2011; Fox, 1995; Mumford, 1995; Trist and Bamforth, 1951). There are however a number of issues when trying to apply current approaches of socio-technical design in a volunteer context:

- These approaches build on upfront planning while it is unlikely that volunteers consider planning and designing their socio-technical infrastructure before starting to work on the project they aim to complete.
- There is a difference between designing CSCW applications, e.g. in research or in an organization compared to volunteer projects, that face a high dropout rate (Kraut et al., 2010).
- Existing approaches on socio-technical design often rely on a common 'background' or 'connection' of the participants. This joint background is created by work contracts, organizational rules and norms which govern collaboration. Being part of an organization thus makes it easier for people to adopt existing work practices, while this is not so likely in the early phases of volunteer collaboration where people are more prone to opt out if they are not satisfied (Haski-Leventhal and Bargal, 2008).
- While it is generally desirable for collaborators to share common interests and values, this is not necessarily the case in an organization. Having common or at least compatible interests and values is however crucial for volunteer projects. Research has shown that maintaining a sense of community is important for sustained volunteer efforts (Cobb et al., 2014).

To create an initial set of categories that can serve as a basis to analyze collaborative practice of volunteers on the fly, we conducted a literature study that focuses on approaches in the context of socio-technical systems (STS), groupware and volunteer collaboration. We also included literature around usability since technology usage will most likely be part of the volunteer's coordination activities. The categories were inspired by the work of Nielsen around usability heuristics since they provide an easy to use set of guidelines to assess complex situations and identify crucial problems (Nielsen, 1994). The literature we studied is spread among but not limited to the following five major clusters:

- socio-technical design (Cherns, 1987; Clegg, 2000; Eason, 2005; Fischer and Herrmann, 2011; Fox, 1995; Mumford, 1995),

- principles of job design (Hackman and Oldham, 1975; Mumford, 1995),
- usability heuristics (Nielsen, 1994),
- principles for the design of computer supported cooperative work and collaboration (Grudin, 1994; Herrmann et al., 1996) and
- volunteer collaboration (Cataldo and Herbsleb, 2008; Chevrier et al., 1994; Cobb et al., 2014; Crowston et al., 2007; Hibbert et al., 2003; Liao et al., 2016).

Category 1 – Reality Check

Category 1 (Cat 1) examines whether the process conducted by volunteers is compatible with the reality of their environment or not. This has been derived from Nielsen's "Match between system and the real world" (Nielsen, 1994).

Suggested questions for detecting problems:

- Is there a sufficient compatibility between pursued goals and what can be achieved in reality based on the available resources?
- Are the terms, information and data being used during the socio-technical process compatible with the language and the information base used by the users of the socio-technical product?
- Are conflicts within the process identified and reported - for example by knowing clients, relevant stakeholders and their interests? This fit with personal interests is related to Mumford's criteria of 'psychological fit' (Mumford, 1995).

Category 2 – Suitability of Task Allocation

Category 2 (Cat 2) is about suitability of task allocation and explores whether tasks are compatible to the competencies and capabilities of volunteers. Clegg refers to the necessity for multiple task allocation (Clegg, 2000); Mumford emphasizes the necessity for a task structure fit (Mumford, 1995).

Suggested questions to detect suitability-related issues:

- Is the distribution of tasks between volunteers and the allocation of tasks understandable and related to volunteer needs, competencies and interests?
- Is there the possibility to assign different arrangements to different volunteers in accordance with their competences, physiological and psychological preconditions and their needs?
- Is coordination communicated to the volunteers and do all volunteers understand it? For example, do volunteers understand how to execute a task, what the available resources are and how the tasks can be carried out efficiently?
- Are volunteers in control of their work (Baker et al., 2001; Herrmann et al., 1996)?

Category 3 – Social Dynamics

Category 3 (Cat 3) explores the role of social dynamics and whether the group accepts and deals with them. Eason claims that socio-technical design has to consider the characteristics of a social system that enables participants in work roles to co-operate effectively (Eason, 2005). In addition, Cataldo and Herbsleb emphasize the necessity for participants to understand their role and how to deal with inevitable fluctuation of volunteers (Cataldo and Herbsleb, 2008).

Suggested questions for identifying problems related to social dynamics:

- Is the relationship clear (described/defined) between the volunteers and the roles they take (e.g. power relations)?
- Is it clear how to deal with ongoing, partially non-anticipatable changes of these relations?
- Is it clear how a volunteer group is prepared to include new volunteers or roles - even if they contribute only informally - or how to deal with changing volunteers?

Category 4 – Proper Information Exchange and Communication Tools

Category 4 (Cat 4) discusses sufficient support and control of information exchange. Cherns includes the necessity for proper information flows (Cherns, 1987); Herrmann et al. require 'suitability of information' (Herrmann et al., 1996); Baker et al. propose principles aiming at proper communication support (Baker et al., 2001). Furthermore, Cat 4 explores the effective integration and efficient use of technology to scaffold communication.

Suggested questions to discover information exchange and communication problems:

- Can volunteers decide - or at least negotiate - which tools they want to use?
- Can volunteers identify what information they need and what information they should provide to the group?
- Are people/roles, who work together, sufficiently connected through spatial conditions, artifacts and communication channels?
- Is sufficient support of communication and information exchange provided and maintained - is this support clearly identifiable?

Category 5 – Balance Between Effort and Benefit, Lack of Motivation

Category 5 (Cat 5) explores the trade-off between the participants' effort and perceived benefits. Grudin mentions related problems in groupware (Grudin, 1994). Values and interests of individuals determine their motivation and willingness of engaging in work (c.f. Cobb et al., 2014; Liao et al., 2016; Mumford, 1995). Similarly, Hibbert et al. (2003) found increased volunteer retention if they perceive their contribution to be worthwhile.

Questions to discover problems:

- Are pursued benefits, goals and the effort how to achieve them clearly described?
- Is it clear how each task will contribute to pursued goals and to values and interests of the volunteers? Are the possible sequences of tasks and workflows clearly directed towards achieving solicited goals/benefits without detours?
- If others are the beneficiaries of one's work: Are the underlying conditions of this exchange clear and transparent?
- Do pursued goals fit motivations and interests of volunteers (Hibbert et al., 2003; Mumford, 1995)?

Category 6 – Feedback and Visibility

Category 6 (Cat 6) deals with providing feedback about outcomes, progress of task completion and options for action. Usability principles in particular emphasize that users must be able to recognize the status of the system, the degree of goal achievement, and have to be guided and supported (Nielsen, 1994). Feedback also is an important aspect of job redesign (Hackman and Oldham, 1975). While visibility in Nielsen's heuristics is an item for evaluating web interfaces, visibility in STS can mean that volunteers show visible motivation to stay in the project or to conduct work. Positive feedback has also been found to increase volunteer retention (Chevrier et al., 1994).

We created following questions:

- Is feedback provided about volunteers' achievements and how well they are acknowledged?
- Is this feedback provided by the coordinator on a substantial basis and at deliberately chosen points in time?
- Do volunteers get guidance according to their needs?

Empirical Method

We analyzed a volunteer group at a mid-western university that jointly developed an AR app over the span of 11 months (09/2015 to 08/2016). Volunteers were from different domains and they were marginally familiar with software and app development. Volunteers did not receive any monetary compensation for their work in the project. A professor, who was interested in augmented reality technology, initiated and continuously supervised the project. This professor sent out a call for participation and ten volunteers responded. None of the volunteers had previous ties with her/him or the department. Out of the ten volunteers who started the project, eight stayed until the end. Two participants dropped out after 6

months (Figure 3). The professor and three researchers (who were not part of the project) conducted the research presented in this paper.

Figure 1. Map view guiding users to points of interest.

Context Information

The goal of the project was to develop an AR app to complement university tours for new students and their families. The idea was to guide users around campus by presenting location-based stories about points of interest (Figure 1). When users arrive at a point of interest, they can use the app to trigger the story (Figure 2). Depending on the spot, users can use different types of media ranging from text to images and video. The app was used during an event to introduce 130 foreign students to the campus.

Project Process

We studied this volunteer project as a socio-technical process in which social settings and technology intertwine (c.f. Figure 3 for an overview of the process of the project). The project started with conceptual meetings (Figure 4 top) during the first month. The volunteers agreed on a preliminary timeline and a meeting schedule, including weekly informal meetings and monthly mandatory meetings between volunteers and the project initiator (Figure 4 bottom). The goal was to set up an initial frame for the project without enforcing a strict project management plan with milestones and deliverables. The group agreed on an initial goal of creating a prototype within eight months (Figure 3). Progress would be discussed during meetings and timelines would be adjusted accordingly.

Figure 2. Game interface at a point of interest on campus.

The volunteers initially aimed at developing the app using GoogleGlass as the main technology. During the first four months, the efforts focused on identifying tools to develop a GoogleGlass app as part of an augmented university tour. It turned out that this would not be possible without major development efforts and monetary funds.

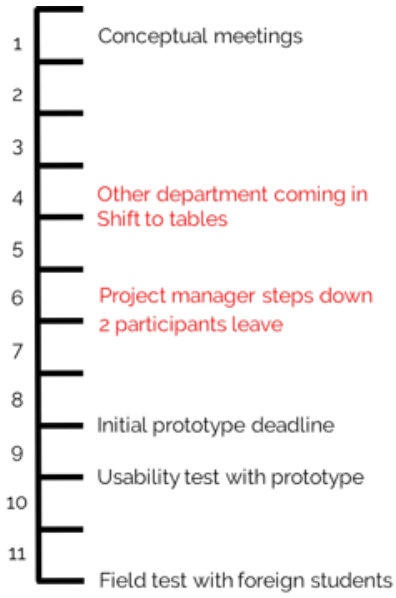

Figure 3. Process of the project over 11 months.

After this investigation, the project took an abrupt change towards using tablets instead of GoogleGlass. This decision was taken as the project initiator was

approached by a university department that was interested in the project. The new stakeholder set a deadline for a field test. This led the volunteers to disband the original plan due to time and resources restrictions and to start working towards a solution based on tablets. At the same time, one volunteer was appointed as project manager.

The volunteers proceeded to search for suitable software and to develop usage scenarios. The scenarios were mainly focused on points of interest around campus. Potential spots were discussed before each volunteer picked a spot and started developing a scenario for it. The scenarios were discussed and refined in follow-up meetings. Around this time, two volunteers dropped out of the project and the project manager stepped down. One month after the initial eight-month deadline, an app was in place and tested with two student groups. Afterwards, changes were made before the app was formally presented and used by a group of 130 foreign students. Results from a study on this large-scale test were mainly positive. The project was thus successful in that the participants developed an app that was perceived as usable and useful by a larger user group.

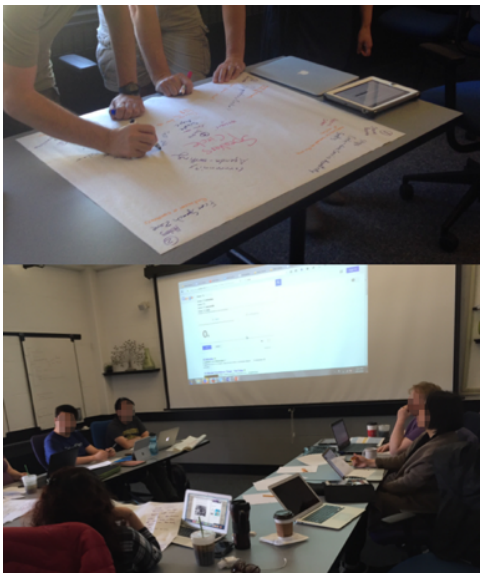

Figure 4. Initial conceptual meetings (top) and monthly project meetings (bottom).

Data Collection and Analysis

Five out of eight volunteers agreed to be interviewed. The others did not respond or they replied they had no time. The interviews lasted between 27 and 57 minutes each. The volunteers covered different career levels (undergraduate, graduate and PhD students, post-docs and faculty), gender (2 female, 3 male), relation to university (no affiliation at all to 25 years of service) and background (information science and education). The interviews were conducted using a semi structured

interview protocol (Denzin, 2008) focusing on collaboration.

To answer the questions stated in the introduction we applied the developed categories to the interviews. Interviews were recorded, transcribed and analyzed using qualitative content analysis (Mayring, 2014). We also had access to documentation from meetings and in between, as well as app versions which served as context during analysis.

Findings

By applying the categories to our empirical case, we identified different episodes during the volunteer group work alongside limitations and ways for improving the categories.

From the interviews, we found multiple indications for discrepancies between project goals and the volunteers perception (**Reality Check (Cat 1)**). The main motive for people to volunteer was an interest to work with GoogleGlass. This is evident by multiple statements such as *"using wearable devices [...] attracted me at the very beginning"* (I1) or *"the technology [...] was innovative and I am highly interested in mixed reality technology"* (I5). Volunteers were also interested in the conceptual idea of the project, to create an augmented learning experience: *"It sounds like this exciting idea conceptually [...] I want to be a part of that"* (I2). Even after it became evident that using GoogleGlass was not feasible (*"without significant funds where we put hundreds of thousands of dollars [...] to try to make it happen"* (I2)), the volunteers were not willing to give up on GoogleGlass until the initiator of the project made *"an arrangement with another department"* (I2). That arrangement required to change to iPads instead of GoogleGlass. The decision that was not unanimously supported: *"I am not so interested in iPad research"* (I4). Our analysis revealed a misalignment between volunteer visions and actual resources for the project. Furthermore, there was a discrepancy between the stories that the volunteers created for campus locations, the length of words and language they used with what would be appealing for future users. This resulted in multiple rounds of *"discussing things over"* (I2) around *"changing the sentences, changing the content, adding more colorful pictures, backgrounds something like that"* (I1).

Volunteers brought diverse skills to the group work. They had backgrounds in *"usability testing"* (I1), *"project management"* (I2) and *"ed[ucation]"* (I5) but they partly over-estimated themselves and their capabilities particularly with respect to their required software development skills. Also tasks were not distributed based on individual skills but based on willingness (**Non-Suitable Tasks (Cat 2)**). This is evident by the statement of a volunteer when asked if s/he could take over a specific task: *"Yeah I can do that"* (I1). Tasks were also rarely coordinated and the volunteers *"failed a lot when tracing [their] activities"* (I3). Realizing that a lack of coordination might become an issue, the group chose one volunteer as a project manager who stepped down after four weeks commenting

that *"this is not a project for project management success"* (I2) and *"it is a very informal group of people who were volunteering"* (I2). During the time where there was no project manager, the group was driven by *"five core members"* (I1). This group, however, did not perceive itself as in charge of coordinating the group but picked up tasks from each other when necessary: *"Some of us ended up redoing [things]"* (I5).

The organization of the project was unanimously described as *"pretty organic"* (I2) and *"self-emerging"* (I3) (**Social Dynamics - Changing Conditions of Organizing the Process (Cat 3)**). There was *"no [explicit] hierarchy"* (I1) and decisions were taken during meetings (*"if we have a meeting, we take a decision"* (I1)). The atmosphere in the project was described as *"very positive"* (I2). The project organization was perceived well by most volunteers, *"the way it happened is probably the way it needed to happen"* (I5). However, there were concerns about the *"lack of structure"* (I5) especially with respect to meetings which were perceived as being inefficient *"that entire situation has made me become very stringent about what meetings I want to attend"* (I2). The same volunteer described the culture of the project as *"meeting happy"* (I2) with *"not much really happen[ing] in these meetings"* (I2). Other volunteers thought that decisions took too long and that the project was *"too slow"* (I4). While most decisions were taken by the group as a whole, few decisions were taken by the project initiator alone. One of these decisions was *"making an arrangement with another department"* (I2) to test the system which led to an abrupt change of plans. Some volunteers also expressed their frustration about the commitment of their peers: *"Not really really interested and motivated"* (I3) which meant that *"some of the [assignments] fell through so some of us ended up kind of redoing those"* (I5). Others perceived their peers as *"driven"* (I2) and *"hard working"* (I2). Not all volunteers appreciated each other in the same way (internal role dynamics). In the beginning of the project, roles were clear and distinct. This changed when personal bonds became stronger. For example, the volunteers referred to the coordinator as a *"kind of friend"* (I1). Other volunteers mentioned that *"it created kind of a friendship"* (I1) or *"there was a camaraderie built because people liked each other"* (I2). Furthermore, the appearance of a new client with new requirements changed the orientation of the project: the project became more formal and less exploratory with a clear goal and delivery deadline (external role dynamics). During the course of the project two volunteers dropped out. One was *"really busy and did not want to develop"* (I3) and the other *"opted out"* (I3). However, despite those two volunteers dropping out we did not find any evidence for a change of dynamics based on those drop-outs.

The volunteers used various ways to communicate, coordinate and distribute content (**Proper Information Exchange, Media (Cat 4)**). The decision for using certain technologies was mainly based on previous experiences and preferences of volunteers, *"s/he is a Box person"* (I2), *"[My colleague] had heard about it"* (I5). This led to tools like GoogleDrive, Box and Samepage all being used at different

points in time for different purposes: from sharing *"notes and schedules"* (I3) to *"storyboards"* (I5) for the app and *"documents"* (I4). None of these tools was used for the entire duration of the project. The usage of different tools at different points in time led to materials being scattered and hard to retrieve. Consequently, volunteers often resorted to using email to communicate: *"[to keep track of what other people are doing] we usually would send emails to ask"* (I4). Email was the only tool used consistently during the whole course of the project: *"email would be number one"*. Nonetheless, email was also criticized for being inefficient (*"too slow"* (I4)). Our analysis thus indicates that control of information exchange, flexibility and autonomy, for volunteers was not sufficiently supported.

As aforementioned, the perception of the effort of other volunteers was not unanimous (**Balance Between Effort and Benefit, Lack of Motivation (Cat 5)**). For example, while one volunteer perceived the others to be *"driven"* (I2) and *"hard working"* (I2), another volunteer stated that people were *"not really really interested and motivated to do this project"* (I3). This may indicate that the core team of project members was motivated and willing to contribute but that was not the case for all volunteers, particularly members who were peripherally associated with the project (*"I am a little on the edge of the project"* (I5)). This assumption is backed up by another project member who states that there were *"core members, like five core members"* (I1). All volunteers described the leadership during the project as very positive: the project initiator was described as being *"great"* (I2), *"extremely hard working"* (I4) and taking *"a lot of effort"* (I1). This motivated some volunteers as evident by the following statement: *"When I saw the project leader is doing great [...] I continued volunteering"* (I1). Some volunteers also described the project as being *"too slow"* (I4) and *"inefficient"* (I4) which led to *"people loose[ing] interest"* (I2). Our analysis indicates different levels of engagement and effort. A difference in effort itself is not problematic. However, it is problematic that this issue had not been addressed during the project. The group did not discuss about different expectations, perceptions and effort. Volunteers may engage differently at different points in time which results in constantly changing conditions of group work compared to non-volunteer project teams. There was an imbalance between what they wanted to achieve, how they expected the others to perform, and the success of the group as whole. The group did not 'see' an appropriate balance between effort and pursued benefits. Still volunteers stayed engaged during the entire project. As a reason for that one member stated that s/he has *"a high work ethic for myself"* (I5). Another person mentioned her/his cultural background as a reason to continue participating in the project: *"if I came in at first and I leave without any good reason [...] it will be considered kind of lazy, not diligent, not hard working, not serious"* (I4). Another reason for people to continuously participate is that people were trying to leave a good impression in order to support their career plans. This becomes evident by one volunteer stating

that *"when you are a student and when you are involved in a research project your reputation is on the line"* (I5).

During the course of the project, there were many opportunities for direct feedback especially due to the fact that there were *"meeting[s] every week"* (I2) which lasted *"for one hour"* (I1) (**Feedback and Visibility (Cat 6)**). The main focus of the meetings however was on *"discussions"* (I1) and *"decisions"* (I3). There was little to no coordination between the members during or in between meetings. In addition, meetings were mainly perceived in a negative way as described under Cat 3. The question thus remains how volunteers stayed motivated. Although feedback was rare, people continued volunteering their time. One reason was that bonds were created between project members (c.f. Cat 3). The results show volunteers' perceived difficulties while conducting the group work. It raises the questions and show problems with respect to visibility, continuous preparation, guidance and an overall supportive environment. The interviews indicate that volunteers did not receive sufficient and explicit feedback about their performance, outcomes, progress of task completion and options for action. Although implicit feedback was provided during regular meetings, the interviews revealed that volunteers did not recognize this as feedback. Nonetheless, it seems that these regular meetings served as a communications channel that contributed to volunteers sticking together and maintaining a common ground, even though this was not the main purpose of the meetings.

Discussion

Applying the categories, we derived from literature to a case of volunteer group work allowed us to develop an understanding of why the volunteers stayed on board and allowed us to detect problems of socio-technical constellations in volunteer group work. Table I provides an overview of issues that could be detected using our six categories.

Table I. Issues discovered from the application of the categories.

Category	Issues detected
Reality Check (Cat 1)	• No sufficient compatibility between pursued goals (goals set by the volunteers) and what could be achieved in reality (the AR Campus tour app). • The language used by volunteers was not always compatible with the language and information requirements of the future users of the system. • Discrepancies between vision and reality during the project were not identified and reported.

Non-Suitable Tasks (Cat 2)	• Tasks did not relate to volunteer competencies. • No possibility to re-assign tasks based on volunteers' competences. • Volunteers understood the coordination of the group even though the process of coordination was not communicated to volunteers. • Guidance was informal and did not meet volunteer's expectations. • The core team did not perceive themselves as in charge of taking decisions.
Social Dynamics (Cat 3)	• No clear role definition for each volunteer and no definition of the relationship between volunteers. • No strategy to deal with ongoing, partially non-anticipative changes of relations between volunteers. • No strategy to deal with changing volunteers roles.
Proper Information Exchange, Media (Cat 4)	• Volunteers had certain preferences for tools based on their previous experiences. These individual preferences were not necessarily compatible. • Volunteers could not decide or negotiate on tool usage. • Volunteers could not identify necessary information and ways to share it with the group. • We assumed that volunteers who worked together, were sufficiently connected to each other by spatial conditions, artifacts and communication channels but the volunteers themselves perceived it differently. • No sufficient support for communication and information exchange was provided.
Balance Between Effort and Benefit, Lack of Motivation (Cat 5)	• The volunteers put a lot of effort, but the perceived benefit was rather low. • Volunteers conducted tasks they did not sign up for. This led to frustration. • There was a gap in the perceived effort between different volunteers. • Volunteers were mainly driven by their intrinsic values. There was little perceived effort for external motivation.
Feedback and Visibility (Cat 6)	• Despite many opportunities, feedback was not provided in a way that it related the volunteer's achievements.

	• Feedback was provided by the coordinator on a substantial basis and at deliberately chosen points in time but not labeled as explicit feedback. • Volunteers mainly worked in solitude on their tasks with little to no feedback.

The categories were thus feasible to detect problems in volunteer collaboration. Nonetheless, they do not directly point towards solutions for them. It can, however, be assumed from the analysis that applying the categories during the course of this project by the volunteers themselves would have supported them in dealing with their problems and it would have probably led to a smoother project process. Volunteer work is a specific context, so it may be that for other kinds of projects additional categories are required or the categories we developed do not apply.

Our analysis helped us understand why people continued volunteering their time despite the unsatisfactory process. The motivation aspect and its relation to personal values is of high relevance and led the volunteers to stick together despite problems they faced during the group work, e.g., organizational issues, unsteady clients and a radical change of technology. From our analysis, we found the following aspects to be the main issues:

- First, for people to get interested initially they need an idea that excites them. In this case, it was the idea of using augmented reality technology (Cat 1).
- Second, the initial motivation may not persist during the course of the entire project. It is thus important to be aware of motivational shifts to not lose volunteers (Cat 2, Cat 6).
- Third, in order to keep people on board for the long run it is necessary to forge relationships among volunteers as well as between volunteers and project leaders. It is necessary to keep track of the social dynamics and remain aware of changes in them (Cat 3).
- Fourth, an important motive for people to stay on board is to further benefit or expand their career potential within an organization. Volunteers should thus have the opportunity to form bonds and create a perspective within the hosting organization (Cat 5).
- Finally, tools can become an additional distraction if they are not well aligned with individual practices or with the organization of the project. They can thus be a source of frustration rather than motivation (Cat 4).

Our analysis also points towards potentials for refining the respective categories so that they are a better fit for volunteer projects:

- **Reality Check (Cat 1)**: While the aspects of this category remain unchanged, it seems necessary for a group to conduct the proposed reality check multiple times throughout a project. Goals and motivations change throughout a project and volunteer groups have to ensure that individuals stick to achievable and interesting goals to retain a high level of commitment.

- **Suitability of Task Allocation (Cat 2)**: There should be a differentiation with respect to the nature of the tasks. Leadership and coordination tasks should be considered as separate tasks that complement practical project work. This became obvious since volunteers seemed to be comfortable to conduct practical tasks. Practical tasks were not well coordinated though because no one did take charge.
- **Social Dynamics (Cat 3)**: This category, similarly to Cat 1, fits the context well. It did however become clear from our analysis that there is a strong inter dependency between tasks (Cat 2) and roles (Cat 3) which should be considered when exploring the social dynamics within a group.
- **Information Exchange and Communication Tools (Cat 4)**: Our analysis indicates that each volunteer brings a set of preferred tools and practices to the group. While different tools can become an issue, it also became clear that the focus should be on the application of tools. Therefore, the focus should be not on the specific tool but on the combination: which tools are used for what and by whom.
- **Balance Between Effort and Benefit, Lack of Motivation (Cat 5)**: The analysis suggested that effort and effectiveness are only parts of a larger picture. It is important that volunteers can decide which tasks they want to take over since they have to perceive them as worthwhile. This requires a certain level of autonomy and control on part of the volunteers which should be taken into account when studying volunteer groups.
- **Feedback and Visibility (Cat 6)**: Bonds between volunteers turned out to be a major factor for their willingness to continue working on the project. This category should thus explicitly include feedback among volunteers in addition to feedback given by the coordinator.

Not all categories were equally important to make sense of the project, its conflicts and problems. The main categories in the studied context of a volunteer group seemed to be those ones that are related to personal interests such as Cat 2 and Cat 5. The categories also pointed towards - sometimes major - problems related to task and collaboration (support), yet the group still stuck together. When there is a clear benefit perceived by the volunteers they stay on the project despite the chaos of the process. It thus seems that not all categories were equally important for this particular project, a point that should be a future subject of study.

Contribution and Limitations

The contribution of this paper is twofold. We propose a set of socio-technical categories based on literature and present results from applying them in a practical context. The application of the categories provides in-depth insights into the socio-technical practice of a small volunteer group that is not part of a larger non-profit organization. This is a subject which has not been studied extensively so far. We

also identified means to improve the categories and discussed their potential application in volunteer group work in general.

Nonetheless, the exploratory nature of this study poses some limitations. First, we drew our initial categories from an analysis of relevant literature. While exhaustive, it is possible that the literature did not cover all aspects that can be found in real world projects. Applying the categories in a project led to deeper insights on their application but it is certainly necessary to confirm their usefulness in further studies. Also, the application of the categories on self-reported data in one project poses a threat to the generalizability of our results. However, our work is meant to be an initial application of newly developed socio-technical categories and thus rather informative than generalizable.

Conclusion and Outlook

This work provides insights into how socio-technical categories can be used to facilitate and to reflect on the collaboration of small volunteer groups outside the context of non-profit or other organizations. The categories can be used by them to evaluate their current practices and identify problems thus leading to a better understanding of volunteer collaboration and improved practices.

In the future we aim to refine the developed categories based on our findings and reflect them on relevant work in the field of co-design (Bratteteig and Wagner, 2014) thus including aspects of power and potential inner-group politics. We then aim to evaluate the refined categories in a larger scale mixed-method study which includes volunteer groups from different domains and analyzes interviews as well as behavioral data and documentation created by the volunteer groups.

Acknowledgments

Dr. Nolte's contributions to this research were partially funded by Deutsche Forschungsgemeinschaft (DFG) under grant no. NO 1302/1-1.

References

Baker, K., S. Greenberg, and C. Gutwin (2001): 'Heuristic evaluation of groupware based on the mechanics of collaboration'. In: *Engineering for human-computer interaction*. Springer, pp. 123–139.

Boulus-Rødje, N. and P. Bjorn (2015): 'Design Challenges in Supporting Distributed Knowledge: An Examination of Organizing Elections'. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. pp. 3137–3146, ACM.

Bratteteig, T. and I. Wagner (2014): *Disentangling participation: power and decision-making in participatory design*. Springer.

Cataldo, M. and J. D. Herbsleb (2008): 'Communication networks in geographically distributed software development'. In: *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. pp. 579–588, ACM.

Cherns, A. (1987): 'Principles of Sociotechnical Design Revisted'. *Human Relations*, vol. 40, no. 3, pp. 153–162.

Chevrier, F., R. Steuer, and J. MacKenzie (1994): 'Factors affecting satisfaction among community--based hospice volunteer visitors'. *American Journal of Hospice and Palliative Medicine*, vol. 11, no. 4, pp. 30–37.

Clary, E. G., M. Snyder, and R. Ridge (1992): 'Volunteers' motivations: A functional strategy for the recruitment, placement, and retention of volunteers'. *Nonprofit Management and Leadership*, vol. 2, no. 4, pp. 333–350.

Clegg, C. W. (2000): 'Sociotechnical principles for system design'. *Applied ergonomics*, vol. 31, no. 5, pp. 463–477.

Cobb, C., T. McCarthy, A. Perkins, A. Bharadwaj, J. Comis, B. Do, and K. Starbird (2014): 'Designing for the deluge: understanding & supporting the distributed, collaborative work of crisis volunteers'. In: *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. pp. 888–899, ACM.

Crowston, K. (2011): 'Lessons from volunteering and free/libre open source software development for the future of work'. In: *Researching the Future in Information Systems*. Springer, pp. 215–229.

Crowston, K., Q. Li, K. Wei, U. Y. Eseryel, and J. Howison (2007): 'Self-organization of teams for free/libre open source software development'. *Information and software technology*, vol. 49, no. 6, pp. 564–575.

Cuskelly, G., T. Taylor, R. Hoye, and S. Darcy (2006): 'Volunteer management practices and volunteer retention: A human resource management approach'. *Sport Management Review*, vol. 9, no. 2, pp. 141–163.

Denzin, N. K. (2008): *Collecting and interpreting qualitative materials*, Vol. 3. Sage.

Eason, K. D. (2005): *Information technology and organisational change*. CRC Press.

Farrell, J. M., M. E. Johnston, and G. D. Twynam (1998): 'Volunteer motivation, satisfaction, and management at an elite sporting competition'. *Journal of sport Management*, vol. 12, no. 4, pp. 288–300.

Farzan, R., R. Kraut, A. Pal, and J. Konstan (2012): 'Socializing volunteers in an online community: a field experiment'. In: *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. pp. 325–334, ACM.

Fischer, G. and T. Herrmann (2011): 'Socio-technical systems: a meta-design perspective'. *International Journal of Sociotechnology and Knowledge Development (IJSKD)*, vol. 3, no. 1, pp. 1–33.

Fox, W. M. (1995): 'Sociotechnical system principles and guidelines: past and present'. *The Journal of Applied Behavioral Science*, vol. 31, no. 1, pp. 91–105.

Garner, J. T. and L. T. Garner (2011): 'Volunteering an opinion: Organizational voice and volunteer retention in nonprofit organizations'. *Nonprofit and Voluntary Sector Quarterly*, vol. 40, no. 5, pp. 813–828.

Grudin, J. (1994): 'Computer-supported cooperative work: history and focus'. *IEEE Computer*, vol. 27, no. 5, pp. 19–26.

Hackman, J. R. and G. R. Oldham (1975): 'Development of the job diagnostic survey.'. *Journal of Applied psychology*, vol. 60, no. 2, pp. 159–170.

Harrison, D. A. (1995): ‘Volunteer motivation and attendance decisions: Competitive theory testing in multiple samples from a homeless shelter.’. *Journal of applied psychology*, vol. 80, no. 3, pp. 371–385.

Haski-Leventhal, D. and D. Bargal (2008): ‘The volunteer stages and transitions model: Organizational socialization of volunteers’. *Human Relations*, vol. 61, no. 1, pp. 67–102.

Herrmann, T., V. Wulf, and A. Hartmann (1996): ‘Requirements for the human centred design of groupware’. *Human Factors in Information Technology*, vol. 12, pp. 77–100.

Hibbert, S., M. Piacentini, and H. Al Dajani (2003): ‘Understanding volunteer motivation for participation in a community-based food cooperative’. *International Journal of Nonprofit and Voluntary Sector Marketing*, vol. 8, no. 1, pp. 30–42.

Karr, L. B. and L. C. Meijs (2006): ‘Sustaining the motivation to volunteer in organizations’. In: *Solidarity and prosocial behavior*. Springer, pp. 157–172.

Kraut, R., M. Burke, J. Riedl, and P. Resnick (2010): ‘Dealing with newcomers’. *Evidencebased Social Design Mining the Social Sciences to Build Online Communities*, vol. 1, pp. 42.

Liao, Q. V., V. Bellotti, and M. Youngblood (2016): ‘Improvising Harmony: Opportunities for Technologies to Support Crowd Orchestration’. *Urbana*, vol. 51, pp. 159–169.

Mayring, P. (2014): *Qualitative content analysis: theoretical foundation, basic procedures and software solution*.

Mumford, E. (1995): *Effective systems design and requirements analysis: the ETHICS approach*. Macmillan.

Nielsen, J. (1994): ‘Enhancing the explanatory power of usability heuristics’. In: *CHI ’94: Proceedings of the SIGCHI conference on Human factors in computing systems*. pp. 152–158, ACM.

Nolte, A. (2018): ‘Exploring Potentials of Process Reflection to Support Communities of Small Volunteer Groups’. In: *Workshops and Work-in-Progress Contributions at S-BPM ONE 2018, Vol. 2074*. CEUR-WS.

Trist, E. and K. Bamforth (1951): ‘Some Social and Psychological Consequences of the Longwall Method’. *Human relations*, vol. 4, pp. 3–3