Chapter 4
Addressing Cooperation Issues in Situated Crowdsourcing

Jorge Goncalves, Simo Hosio, Niels van Berkel and Simon Klakegg

Abstract Situated crowdsourcing has been growing in popularity as an alternative way to collect complex and often creative crowd work. However, previous situated crowdsourcing deployments have not successfully leveraged cooperation possibilities with their audiences, which can improve the data quality of deployed macrotasks. In this chapter, we present three situated crowdsourcing case studies that used different situated technologies and identify the reasons behind their missteps regarding promoting cooperation between workers. Then, based on the identified issues, we propose the design of a novel situated crowdsourcing platform that aims to effectively support cooperation without alienating solo workers. In order to gather insights on our proposed design, we built a prototype platform and evaluated it using a laboratory study with 24 participants. In general, participants were positive about the idea as it provided an easy way to cooperate with friends when completing tasks, while also allowing them to adjust the working environment to their liking. Finally, we conclude by offering insights towards improving cooperation in future situated crowdsourcing deployments and how this can assist in completing macrotasks.

4.1 Introduction

Situated crowdsourcing has emerged as a promising new crowdsourcing paradigm, aimed at providing a complementary means to elicit crowd contributions (Hosio et al. 2014). It entails embedding situated input technologies (e.g. public displays, tablets) in a physical space and leveraging users' serendipitous availability (Müller et al. 2010) or idle time [‘cognitive surplus’ (Shirky 2010)]. Due to its' characteristics, situated crowdsourcing enables the collection of crowd contributions that can be

J. Goncalves (✉) · N. van Berkel
School of Computing and Information Systems, The University of Melbourne, Parkville, VIC 3010, Australia
e-mail: jorge.goncalves@unimelb.edu.au

S. Hosio · S. Klakegg
Center for Ubiquitous Computing, University of Oulu, Pentti Kaiteran Katu 1, P.O. Box 4500, 90014 Oulu, Finland

V.-J. Khan et al. (eds.), *Macrotask Crowdsourcing*,
Human–Computer Interaction Series, https://doi.org/10.1007/978-3-030-12334-5_4

challenging to gather with other forms of crowdsourcing (e.g. online). For instance, it allows for targeting of specific individuals' in a certain location (Goncalves et al. 2014a; Heimerl et al. 2012), gathering people's local knowledge on a particular topic (Goncalves et al. 2014b) or reaching an untapped source of potential workers (Hosio et al. 2014). For these reasons, the number of situated crowdsourcing deployments reported in literature is on the rise (e.g. Heimerl et al. 2012; Goncalves et al. 2013, 2016, 2017; Hosio et al. 2014; Huang 2015; Ludwig et al. 2017).

Situated crowdsourcing also opens up opportunities to conduct macrotasks by targeting workers with specific expertise or knowledge. However, macrotasking often involves worker cooperation (Schmitz and Lykourentzou 2018), which is an important challenge with situated crowdsourcing deployments as it can be difficult to promote and/or design for cooperation between the workers. One of the reasons behind this is the inherently public nature of the situated technologies used in these deployments, which has been shown to sometimes lead to disruptive and non-serious behaviours (Kuikkaniemi et al. 2011). Further, while online and mobile crowdsourcing allows each individual to use their own personal device, and facilitate the design of tasks that support cooperation, in situated crowdsourcing there is typically only access to one single device within a specific location. At the same time, situated crowdsourcing deployments in literature have not provided appropriate scaffolding to support cooperation between users, which further exacerbates the aforementioned issues.

In this chapter, we summarise the findings of three situated crowdsourcing deployments using different types of situated technologies (single-purpose large public displays, crowdsourcing kiosks embedded with tablets, and multipurpose public displays) in terms of cooperation between the workers, and discuss the lessons learned. We then propose a novel design for a situated crowdsourcing platform to better support cooperation between workers based on these lessons, which in turn can facilitate the completion of macrotasks. Finally, we present preliminary qualitative results on users' opinions on the prototype design's appropriateness for crowdsourcing and discuss the potential of situated crowdsourcing with regards to the deployment of macrotasks.

4.2 Related Work

4.2.1 Cooperation in Crowdsourcing

Online crowdsourcing platforms have enabled cooperation between workers using computational systems without any limiting spatiotemporal boundaries. However, cooperation is rarely an explicit feature of the work. It is the requesters who must divide, distribute and combine the received work to make it a cohesive whole (Martin et al. 2016). For instance, using *Etherpad* (a lightweight collaborative online notepad), workers from MTurk have been successfully tasked with translating Spanish poems into English (Kittur 2010). *Flash Teams* is a framework to coordinate experts from a crowd to perform, e.g. rapid design prototyping or course development (Retelny et al. 2014). They also explore how to create entire organisations consisting of teams with different skill sets which can practically provide output 24 h per day, as the workforce is truly global. Yet another example is *Huddler*, that is used to assemble familiar teams during uncertain availability from MTurk. Huddler (Salehi et al. 2017) provides a thin wrapper where workers wait for other workers to join the ad hoc team before proceeding to complete the actual tasks. Haas et al. presented Argonaut, a framework that improves macrotask-powered work quality using a hierarchical review (Haas et al. 2015).

However, crowdsourcing platforms do not always support cooperation between its users and there is great variation in the extent and nature of the collaboration that occurs (Saxton et al. 2013). Innocentive is a good example of a crowdsourcing platform that only permits partial collaboration in order to safeguard the intellectual property of the task requesters. When its users receive notification about available challenges, they can either tackle it as an individual or with agreed-upon team members available through a confidential Team Project room. Similarly, workers of the platform Upwork can either complete the tasks alone or invite other community members to form a project group.

Beyond cooperation within crowdsourcing platforms, workers of Amazon's Mechanical Turk (i.e. Turkers) have developed elaborate ways to cooperate to identify lucrative tasks or to recreate the social aspects that exist in traditional brick and mortar work (Gray et al. 2016). This is important, as it has been argued that requesters can in some cases benefit from the lack of cooperation amongst workers (Felstiner 2012). Hence, external tools that support cooperation amongst Turkers so they can work together to exert more control over the crowdsourcing market (Martin et al. 2014) are sometimes necessary. Ultimately, Turkers' influence over the platform will depend on the available tools and on how much workers are willing to share their perspectives and actions with others (Martin et al. 2014). Furthermore, a better understanding of how work is actually done can help designers and software engineers who are developing tools to support that work (Gupta et al. 2014). Here, we highlight the challenges of cooperation in situated crowdsourcing and propose a design aimed at providing appropriate scaffolding for cooperation without relying on external tools.

4.2.2 Situated Technologies and Their Use for Crowdsourcing

A key human characteristic that situated technologies, such as displays, can exploit is the fundamental need to explore, to start using pieces of technology rather serendipitously and simply to 'kill time' (Müller et al. 2010). Thus, situating the deployment somewhere people typically have free time is considered beneficial. Furthermore, the presence of users who are publicly interacting with a deployment draws the attention of passers-by—a phenomenon better known as the *honey-pot effect*—as observed by Brignull and Rogers (2003). The honey-pot effect can be leveraged to increase interactions with a deployment simply by designing for attention and affording the audience to start using the deployment (Hosio et al. 2016). However, deployments often tend to support only one simultaneous user, and the honey-pot effect leads to queuing, which can be detrimental to the overall experience.

Furthermore, situated technology deployments are often used by groups of users (Hosio et al. 2016). However, while using technologies in groups of people is fun and entices interaction, social awkwardness has also been documented in such situations. For example, the space around a public deployment can be perceived as a proverbial stage where the audience is watching the user (Kuikkaniemi et al. 2011). The group members also sometimes conflict with each other when using a shared deployment. For example, Peltonen et al. studied social group interactions in an urban city area using a public display as an intervention (Peltonen et al. 2008). They document in detail how the presence of users invited new interactions to take place with the deployment, and how the presence of others often leads to conflicts and tensions in the personal spaces of users.

Situated technologies have certain desired characteristics for crowdsourcing, such as low barrier of entry for people who would not otherwise engage in crowdsourcing or targeting a specific group of wanted participants (Goncalves et al. 2013; Hosio et al. 2014). However, the aforementioned issues with situated technologies also hinder their potential for crowdsourcing purposes, and with our work, we seek to pinpoint and offer solutions to identified cooperation challenges.

4.2.2.1 Situated Crowdsourcing Deployments

Crowdsourcing using situated technologies is becoming more and more feasible as the number of installations grows. A recent example of a situated crowdsourcing deployment is *Umati*, an augmented vending machine used to explore CommunitySourcing (Heimerl et al. 2012). Umati dispatched edible goods such as snacks and chocolate in exchange for labour that could only be completed accurately by local workers. *Bazaar*, by Hosio et al., investigated how an economic market model applies in situated settings, concluding that the supply of labour can indeed be controlled with alternating the rewards also in situated task markets (Hosio et al. 2014). The same platform was later used to explore the collection of subjective and local data as well (Goncalves et al. 2017). Two more recent examples include *CrowdFeedBack* and *CrowdButton* that together focus on sustaining the uptake and quality of unpaid crowdsourcing contributions (Huang 2015). As a final example, *City-Share* facilitates efficient communication between official emergency personnel and volunteers in disaster zones by using public displays as communication hubs (Ludwig et al. 2017). With the rise of situated crowdsourcing deployments, Huang et al. (2017) proposed a genetic model inspired by the MIT's model on collective intelligence (Malone et al.

2010), aimed at helping researchers in this area by identifying important contextual aspects for user contributions in situated crowdsourcing systems.

Despite the much-explored potential, situated crowdsourcing deployments are inherently limited by both scale and reach. Contrary to traditional online crowdsourcing, where a deployment can potentially reach millions of users (Ipeirotis and Gabrilovich 2014) who contribute using their own familiar devices anywhere, in situated settings the workers typically complete tasks using devices deployed by third parties as parts of the fixed environment. For this reason, researchers consider situated crowdsourcing more as an alternative, or different means of eliciting crowd contributions, rather than a replacement or competitor of online crowdsourcing (Goncalves et al. 2013).

4.3 Case Studies

Next, we summarise the findings regarding cooperation between workers of three separate situated crowdsourcing deployments that used different types of situated technologies, namely: (1) single-purpose public displays, (2) multipurpose public displays and (3) kiosks embedded with tablets.

4.3.1 Case Study 1 (C1): Crowdsourcing Malaria Detection

Our first case study entailed using four 46″ single-purpose public displays (Fig. 4.1) to crowdsource malaria detection. The task entailed asking workers to count malaria-infected blood cells on images of a petri dish generated algorithmically while comparing different motivational approaches. More details on this deployment can be seen in Goncalves et al. (2013).

What makes this particular deployment unique is that we recorded all interactions with one public display, and thus we were able to observe participants' attitudes and social context when completing tasks. Our video recordings consisted of 123 distinct instances of interaction and based on content analysis using open and axial coding we identified different emerging themes of behaviour. As reported in Goncalves et al. (2013), this analysis confirmed instances of the behaviours that we initially noted in our in situ observations, but also revealed several new behaviours that people exhibited when using the display. The six identified behaviours were:

- **Ignorer**: passers-by that ignored the display, exhibiting what is often referred to as display blindness (Müller et al. 2009), and
- **Unlocker**: those that actually unlocked the screen but completed no tasks. These account for the high number of curiosity clicks mentioned previously.
- **Herder**: individuals would approach the display with a group of people, complete some tasks and then leave with the group. The other members would adopt a

Fig. 4.1 Example of one of the single-purpose public displays used in this deployment

passive position behind the herder, in a way that suggested they were not applying social pressure but rather observing,

- **Loner**: individuals that approached the display alone and typically spent more time than others completing tasks.
- **Attractor**: attracted others to join them on the display, commonly referred to as the honey-pot effect (Brignull and Rogers 2003), and complete tasks jointly.
- **Repeller**: applied social pressure to try to make the worker leave the display. Instances of repellers also happened when groups of two or more people approached the display.

A visual representation of each of these behaviours can be seen in Fig. 4.2. Overall, analysis of the work conducted by each group of workers showed that sole users, dubbed as *loners*, spent more time completing tasks. More specifically, loners completed on average a higher number of tasks ($M = 4.91$) when compared to the other groups: attractors ($M = 3.71$), herders ($M = 3.43$) and finally repellers ($M = 1.29$). A Kruskal–Wallis test showed that there was a significant difference in average number of tasks completed between the different behaviours ($\chi 2(4) = 22.18$, $p < 0.01$). Post hoc analysis using the Mann–Whitney tests showed that there was only a significant difference between loners and repellers in terms of average number of completed tasks ($U = 26.04$, $p < 0.01$). As for accuracy, a Kruskal–Wallis test showed that there was no significant difference in accuracy between the different behaviours ($\chi 2(4) =$

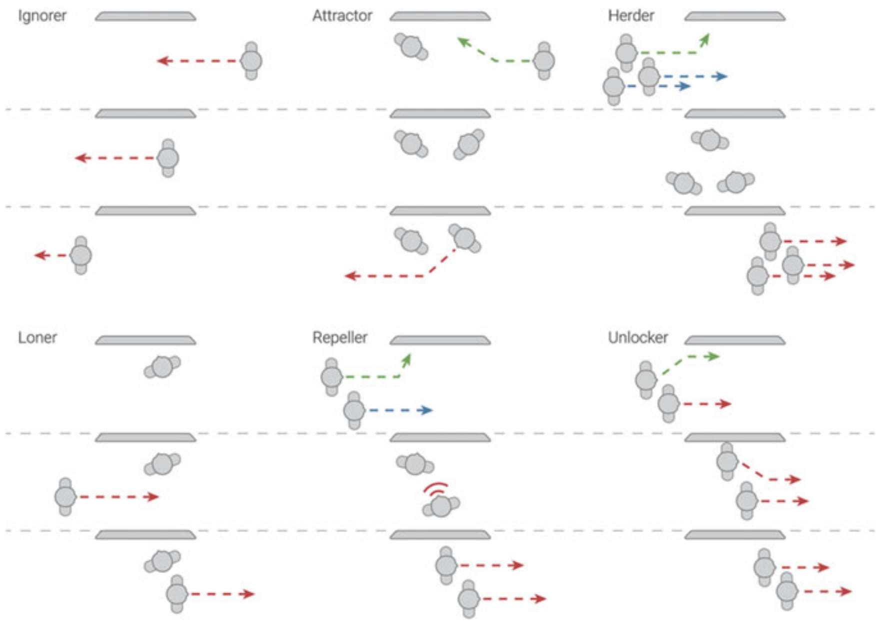

Fig. 4.2 The six identified behaviours in this deployment

7.99, $p = 0.09$). These results suggest that while not having a significant impact on accuracy, groups of workers spent significantly less time completing the tasks.

These results can be seen as problematic as situated technologies naturally invite groups of people to engage with them or have people join those already engaging with the technology during interaction [known as the honey-pot effect (Brignull and Rogers 2003)]. Furthermore, previous work has suggested that this behaviour is effective in combating feelings of self-consciousness felt by a solo user when engaging with public technologies (Kuikkaniemi et al. 2011). Unfortunately, the latter two behaviours (attractor and repeller) ultimately led to a disturbance and delay in the completion of the tasks. Here, the workers were not encouraged to perform well, but instead engage in performative acts (Hosio et al. 2015) resulting in non-serious completion of tasks. In fact, previous work suggests that in some cases the engagement with these interactive public artefacts emerges only when the overall social context provides a 'license to play' (Jurmu et al. 2014). In the case of playful applications or games, this does not matter and can even act as a catalyst to use (Kuikkaniemi et al. 2011), but for crowdsourcing purposes where meaningful data is being collected from the public, it is important to provide appropriate scaffolding for group use. If additional individuals feel like they are not able to contribute meaningfully, then this will ultimately lead to them disturbing those that are engaging with the platform. While this particular case study involved microtasks, we argue that our findings generalise to the completion of macrotasks in a situated crowdsourcing deployment (i.e. similar context).

4.3.2 Case Study 2 (C2): Crowdsourcing Public Opinion

Our second case study deals with a public, large-scale in-the-wild deployment at the heart of downtown Oulu, in Finland. Collecting and analysing city-scale feedback from individual citizens is one way of crowdsourcing the public opinion (Hosio et al. 2015). In this case study, we used a grid of interactive large public displays, *UBI-hotspots* (Hosio et al. 2016), to elicit civic feedback from the young (Hosio et al. 2015). More specifically, we deployed a photo booth application that was paired with social media to enable a two-way communication channel between citizens and officials.

UBI-hotspots (as seen in Fig. 4.3) are large displays deployed in pivotal locations in Oulu. These displays host several applications, i.e. they are *multipurpose* (Hosio et al. 2013). This ‘battle’, where every application has several contenders for user attention, led to designing the application as playful in the first place. At the time, we reasoned it is fair to anticipate the younger generations to be drawn into gamified concepts rather than ‘boring’ civic affairs. The key design choices in addition to playfulness were to exploit the attractiveness of public technologies in general (Müller et al. 2010) and to extend interaction capabilities by leveraging social media.

In terms of the original goal, i.e. providing a useful two-way discussion channel between the young and the city youth affairs department, the six months-long deployment turned out to be a quite the fiasco. While the volume of submissions, or feedback items, was fairly satisfactory (425 unique submissions), it soon became painfully clear that practically none of them had anything to do with the original goal of the deployment. No feedback was being crowdsourced, and the deployed system was used for toying around and for taking snapshots for the sake of having fun. A representative sample of the submitted entries can be seen in Fig. 4.4.

In hindsight and regards to situated crowdsourcing, we identify an important aspect worth considering in the design stage. While playful design elements that are often praised in related literature as good ways to elicit engagement, it backfired in

Fig. 4.3 Example of one of the multipurpose public displays used in this deployment

Fig. 4.4 Citizen feedback submitted through our crowdsourcing platform. Top row: teenagers playing with an energy drink can, tourists taking pictures with the deployment, teens acting to the camera. Bottom row: groups of people posing for the camera

this type of 'serious' application. Granted, the young did enjoy using the application, and at times spent several minutes with it in order to create beautiful sequences of pictures, but for the 'wrong' purpose. One can learn a lot from human behaviour such as demonstrated in the submissions (Hosio et al. 2015), but this takes a lot of effort and does not necessarily answer to the original needs of the deployment. Providing feedback to the city was simply a lesser motive than having fun with the tech just to take 'funny' pictures. That being said, the big screens we used were clearly suitable for ad hoc cooperation to take place: the large screens were used as toys to play with, and especially the camera was seen as a motivator to engage with the application. In that sense, designing for playfulness that channels the energy and exploration to the intended direction can be beneficial.

4.3.3 Case Study 3 (C3): Situated Crowdsourcing Market

Our third and final case study entailed the development and deployment of a situated crowdsourcing market, called *Bazaar*, using multiple public kiosks embedded with tablets deployed in different locations (Fig. 4.5). The platform enabled users to create accounts, earn virtual currency by completing a number of different types of tasks (e.g. sentiment analysis, image labelling) and exchange earned currency for rewards (e.g. money, movie tickets, coffee vouchers, etc.). More details on this deployment can be seen in Hosio et al. (2014).

Here, one of our intentions was to provide a more private means to complete crowd work and mitigate any self-conscious issues when engaging with public technology. However, as a direct result of the smaller screen estate, collaborative work between

Fig. 4.5 Example of one of the kiosks used during this deployment

the workers became more challenging. In fact, during our interviews, several users of *Bazaar* reported wanting to work towards a common prize or simply help a friend complete the given tasks. Unfortunately, the platform did not support this, which in some cases resulted in the workers quitting the platform altogether. Those with friends that continued using the platform found alternative ways to achieve their goals, such as sharing accounts or, more commonly, working separately in different locations instead of cooperating in a meaningful way. Given the distance between the different kiosks, this solution proved to be rather non-ideal removing any social aspects from conducting the crowdsourcing work. Several groups of workers completed tasks until they all achieved a certain goal (i.e. each person getting enough virtual currency to get a movie ticket), and then stopped using the platform. In case a similar deployment was to be conducted in the future to support cooperation between workers, then a collocated solution could prove more efficient in attracting and engaging workers with the platform. Finally, the design was deemed as not enough customizable in terms of ergonomic factors: workers wanted to adjust the height or even the angle of the display, as in many cases the sun or other lights were reflecting from the embedded tablet's surface.

4.3.4 Summary of Identified Issues

First, one major pitfall in our presented case studies (and other situated crowdsourcing deployments), is that they did not allow more than one person to directly engage with the tasks simultaneously. This can ultimately lead to a disturbance that will affect the worker engaging with the tasks (as seen in C1 and C2). One potential solution is to allow additional people to use their own personal devices to contribute to the crowdsourcing task (e.g. mobile phone). However, it is challenging to provide reliable runtime assembly of multi-device ecologies (Heikkinen et al. 2014;

Weißker et al. 2016), and without seamless interactivity, workers can quickly lose interest. Furthermore, previous work has shown that adding additional barriers to participation can significantly hinder the likelihood people will engage with a situated crowdsourcing platform (Goncalves et al. 2013). Hence, offering a simple and rapid solution to enable cooperation in these settings is crucial.

Second, the type of situated technology will significantly affect what kind of work can be conducted and how cooperation should be supported. Given the added control, better usability and more private crowdsourcing experience (i.e. smaller screens meant that others could not see what a worker was doing) of situated kiosks, we argue that they are better suited to support cooperation in situated crowdsourcing deployments. However, while the experiment reported in C3 did indeed offer these benefits, it also restricted even further any possibility for cooperation. Several workers that interacted with the platform were eager to cooperate with others, and ended up taking alternative routes to achieve this goal. Thus, we argue that a multiple input and collocated solution would trump multi-location deployments (such as the one reported in C3) when cooperation between workers is desired.

Finally, while designing with playfulness in mind has been showed in the past to be highly successful in engaging users with situated technologies (Kuikkaniemi et al. 2011), in C2 it was highly detrimental to the original intent of the experiment: to crowdsource public opinion on a specific matter. This is not to say that performing tasks in a situated crowdsourcing environment should not be enjoyable, but that the design should minimise as much as possible appropriation of the technology by workers for different purposes than originally intended.

4.4 Proposed Design

In order to mitigate the issues identified in our case studies, we designed and constructed a situated crowdsourcing table with three attached tablets (Fig. 4.6). The design of this table was informed by the findings reported in our case studies, as well as years of experience conducting situated crowdsourcing experiments. We settled for three tablets as we rarely saw larger groups engage with the display in C1 and also because it allows for a few to few ecosystems that enables natural interaction between the workers to occur (Terrenghi et al. 2009). This is not to say that the platform would enforce three simultaneous workers, but allow for up to thrrree workers to interact with the available crowdsourcing tasks. The proposed design also enables solo workers to complete tasks if they so choose, including simultaneous solo workers that do not wish to cooperate. It would then be up to the task requester and/or designer to decide which tasks available on the platform would support cooperation and which would not. When designing for cooperation this could be achieved directly on the interface (e.g. workers interact with the same task simultaneously to solve it), or indirectly by simply encouraging communication between the workers (e.g. each worker interacts with different subtasks of a larger task). For instances of direct cooperation, assigning a leader may be necessary to ensure high-quality task

Fig. 4.6 Situated crowdsourcing table with three attached tablets

completion, and to coordinate and submit the work, as suggested in Retelny et al. (2014). We also anticipate having a responsive leader in each session to, at least in some cases, reduce the amount of non-serious behaviour in other workers.

Furthermore, the tablets are placed within a special enclosure to prevent appropriation of the technology as seen in C2 (e.g. power button is inaccessible). Furthermore, a registration process required before completing any tasks can filter out non-serious individuals (Hosio et al. 2014). The enclosure rests upon a hinge, allowing workers to reposition the tablet vertically [adapting the visual angle as suggested in Terrenghi et al. (2009)] and potentially show to the other workers what is currently on their screen. In addition, the enclosure allows the worker to rotate the tablet as deemed necessary. We opted for a round table to promote conversation and cooperation between workers currently working on the same task, as seen in Shen et al. (2003). The table's height is also adjustable to cater to a more diverse set of potential workers and promote inclusivity. While there will be issues when workers of very different heights engage with the platform, we argue that this design is still more inclusive than past situated crowdsourcing deployments reported in the literature that uses a technology with a predefined and non-changeable height. As an example, the kiosks presented in C3 did not allow workers to adjust the height of the screen or the visual angle, making for a non-ideal working experience for some workers. A summary of the identified issues and the design choices aimed at solving them can be seen in Table 4.1.

4.4.1 Interview Procedure and Method

We recruited 24 participants from mailing lists in our university and social media. Recruited participants were from several different study areas such as computer science, biomedical engineering, biology, education, and product management. In our usability lab, we showed the participants the table and allowed them to directly

Table 4.1 Identified issues from our case studies and solutions offered by our proposed design

Identified issue	Design choices
Idle friends disrupting others' work (C1, C2)	Multiple collocated devices
Excessive appropriation (C2)	Enclosure that hides certain functions. Appointing a responsible leader, requiring registration
Cooperation not supported (C1, C3)	Generic platform that allows, but does not enforce cooperation in tasks
Physical limitations, work ergonomy (C3)	Adjustable table design (height, device angle, rotates)

interact with it. We conducted the semi-structured interviews in groups of 3 with each one lasting around 15 min. During the interviews, we asked their opinions regarding the design of the table, what tasks would work well or not with the setup, and the benefits and drawbacks the proposed design would have over other situated technologies (e.g. large public display) for completing macrotasks. Participants were given a movie voucher for their participation.

We used thematic analysis to explore our qualitative data. Thematic analysis is 'a method for identifying, analysing, and reporting patterns (themes) within data' and is commonly applied in qualitative research (Braun and Clarke 2006). First, we extracted the qualitative data from our responses, and focused on discovering different themes. We then wrote simple descriptive notes on these themes and discussed them. Since our research is largely exploratory without a theoretical framework about designing for cooperation in situated crowdsourcing, our coding process was inductive. Codes emerged and were selected through an iterative process and discussion between the coders.

4.4.2 Results

4.4.2.1 Input Mechanisms

Participants expressed that the number of available tablets would most likely be sufficient in most cases, but at the same time could see how a higher number of tablets could sometimes be useful. By offering several simultaneous input mechanisms it is more likely that present individuals express their opinions when compared to the typical single input mechanism platform reported in the majority of situated crowdsourcing deployments. This was seen as particularly useful in the case of macrotasks where simultaneous input could facilitate the completion of the tasks. This effectively breaks these tasks into microtasks, which has been shown to result in higher quality outcomes and a better experience that can reduce the impact of interruptions (Cheng et al. 2015).

> If this happens on one screen—maybe some people might not express their opinion. So if we have three tablets we can be sure that everyone mentions their opinion. (P12)

> Definitely with more complex tasks having separate inputs is great, instead of everyone trying to chime in on the same screen. Less confusion and more likely that everyone contributes. (P17)

4.4.2.2 Table Design

Furthermore, one group of participants appreciated the privacy aspects of our design, stating that it would be much more awkward to complete tasks on a larger display. This is in line with previous work on public displays that report feelings of self-consciousness when interacting with a large display in public areas (Kuikkaniemi et al. 2011).

> I would feel awkward or embarrassed when doing it on a larger screen, so I prefer smaller screens for this. (P02)

This is of particular importance when completing more sensitive tasks that workers might, in general, be less comfortable completing, and may even prefer completing them alone, a possibility that is also possible in our proposed design.

In addition, several participants identified the repositioning features of the tablet enclosures as a beneficial way to quickly show others what is on their screen, thus supporting cooperation between the workers. Finally, participants appreciated the ability to rotate the tablet to a more comfortable position as typically situated crowdsourcing deployments can be quite tiring when completing tasks for an extended period of time.

4.4.2.3 Collocated Interaction

While all situated crowdsourcing deployments have elements of collocated interaction, participants reported that our design could further facilitate these interactions. The round design of the table and closeness of each tablet was seen as an important enabler for better communication between the workers. Unlike situated crowdsourcing deployments that use public displays and have workers stand side by side, our design positions workers to be face-to-face facilitating interaction.

> Communication is better, you can see the faces, impressions, and everything. (P05)

> I like the fact that it is a round table, because you can see each other faces. It facilitates conversation, a big screen would be worse. You cannot experience the feelings of people etcetera. I also kind of like that everyone has their own screen. (P08)

In addition, workers can more easily identify if others are unsure of their answers or not contributing sufficiently to the tasks.

> It can help as you can see the body language or if someone is a little bit shy or not saying things. (P22)

4.4.2.4 Task Suitability

During the interviews, some participants expressed the suitability of different types of tasks to the proposed design. For instance, visual search tasks (e.g. finding a certain object in an image) would benefit from all workers interacting on the same screen.

> If we could have one big screen for searching, that would be good. Just one screen for all of us. (P13)
>
> If we all have one big screen, it is easier to see what everyone is looking at - or are pointing. (P20)

This can be explained by the fact that such tasks are objective and have only 1 correct answer. By having all participants look at a single screen will lead to faster completion times, and therefore a more efficient workflow. However, for cooperation in most types of tasks, participants agreed that the proposed design would be more advantageous over a larger public display. For instance, in more subjective tasks workers are able to discuss and potentially annotate parts of a task without disturbing the view of the others.

4.4.3 Lessons Learned

In this section, we summarise the lessons learned through the design and evaluation of our situated crowdsourcing platform. Situated crowdsourcing enables crowd work that requires local knowledge or that benefits from face-to-face interactions, tasks that are challenging to complete with online crowdsourcing, so appropriately supporting this collaboration is crucial. Participants of our user study praised the approach for allowing easy collocated cooperation between workers and adjusting the work environment to their specifications. In addition, the use of tablets over large public displays was mostly seen as beneficial in preserving privacy as well as promoting discussion between the different workers. Moreover, the better usability of these devices can facilitate the completion of macrotasks, which can be challenging to complete in a situated crowdsourcing setting due to task complexity and the increased likelihood that workers can be distracted by the surrounding environment. Furthermore, while the proposed design may not be ideal for cooperation in every type of tasks, it was considered as being an effective approach to provide, in most cases, sufficient scaffolding for cooperation between situated crowdsourcing workers. Finally, while completing macrotasks using our design is likely to result in longer completion times, it is also likely to result in higher quality outcomes and a better experience as it breaks these tasks into more manageable microtasks (Cheng et al. 2015).

In general, it is crucial for researchers to conceptualise new forms of crowd work that go beyond simple and independent tasks that are common today in many crowdsourcing platforms (Kittur et al. 2013). In the case of situated crowdsourcing, allowing and supporting cooperation between collocated workers presents itself as an important research direction, as macrotasking often involves worker cooperation (Schmitz and Lykourentzou 2018). In that sense, our design was considered by our participants as a positive step towards effective cooperation in situated crowdsourcing settings, as it has the necessary characteristics to facilitate conducting work in a more challenging setting when compared to online crowdsourcing.

4.5 Conclusion and Future Work

Previous deployments in situated crowdsourcing leveraged little or no cooperation between the anticipated workers, thus making it challenging to deploy complex tasks. We argue that this is caused not by an inherent limitation of situated crowdsourcing, but instead it is due to the fact that these deployments did not provide appropriate scaffolding to support said cooperation. With this chapter, we identify specific challenges and flaws in design that have led to this potential shortcoming, in order to inform researchers interested in conducting situated crowdsourcing experiments. Namely, lack of support for several simultaneous workers, inefficient distribution of input mechanisms, design that allowed for appropriation, among others were identified as important challenges that should be considered when designing situated crowdsourcing experiments that support cooperation between workers. Taking these identified challenges into consideration, we then proposed our own design of a situated crowdsourcing platform that facilitates cooperation between workers, and therefore, the completion of relevant macrotasks.

In the future, we hope to implement and evaluate a situated crowdsourcing market that leverages the table design proposed in this chapter. This would entail designing different crowdsourcing tasks for both solo and groups of workers, and conducting an in-the-wild deployment. Ultimately, we argue that it is important to develop new situated crowdsourcing ecologies that support, not enforce, cooperation between workers engaging with the platform and we believe our work presents an important first step towards this goal.

Acknowledgements This work is partially funded by the Academy of Finland (286386-CPDSS, 285459-iSCIENCE), the European Commission (Grant 6AIKA-A71143-AKAI), and Marie Skłodowska-Curie Actions (645706-GRAGE).

References

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Brignull, H., & Rogers, Y. (2003). Enticing people to interact with large public displays in public spaces. In M. Rauterberg, M. Menozzi, & J. Wesson (Eds.), *Proceedings of 9th IFIP TC13 International Conference on Human-Computer Interaction, INTERACT '03, Zurich, Switzerland, 1–5 September 2003* (pp. 17–24). Amsterdam: IOS Press.

Cheng, J., Teevan, J., Iqbal, S. T., & Bernstein, M. S. (2015). Break it down: A comparison of macro- and microtasks. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)* (pp. 4061–4064). ACM Press.

Felstiner, A. (2012). The weakness of crowds. *Crowds and Clouds.* Retrieved March 15, 2017, from http://limn.it/the-weakness-of-crowds/.

Goncalves, J., Ferreira, D., Hosio, S., Liu, Y., Rogstadius, J., Kukka, H., et al. (2013). Crowdsourcing on the spot: Altruistic use of public displays, feasibility, performance, and behaviours. In J. F. Canny, M. Langheinrich, & J. Rekimoto (Eds.), *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '13,* Zurich, Switzerland, 8–12 September 2013 (pp. 753–762). New York: ACM Press.

Goncalves, J., Pandab, P., Ferreira, D., Ghahramani, M., Zhao, G., & Kostakos, V. (2014a). Projective testing of diurnal collective emotion. In J. Kientz, J. Scott, & J. Song (Eds.), *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '14,* Seattle, USA, 13–17 September 2014 (pp. 487–497). New York: ACM Press.

Goncalves, J., Hosio, S., Ferreira, D., & Kostakos, V. (2014b). Game of words: Tagging places through crowdsourcing on public displays. In C. Neustaedter, S. Bardzell, & E. Paulos (Eds.), *Proceedings of the 2014 Conference on Designing Interactive Systems, DIS '14*, Vancouver, Canada, 21–25 June 2014 (pp. 705–714). New York: ACM Press.

Goncalves, J., Kukka, H., Sanchez, I., & Kostakos, V. (2016). Crowdsourcing queue estimations in situ. In P. Bjørn & J. Konstan (Eds.), *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work and Social Computing, CSCW '16,* San Francisco, USA, 27 February–2 March 2016 (pp. 1040–1051). New York: ACM Press.

Goncalves, J., Hosio, S., & Kostakos, V. (2017). Eliciting structured knowledge from situated crowd markets. *ACM Transactions on Internet Technology, 17*(2), Article 14.

Gray, M. L., Suri, S., Ali, S. S., & Kulkarni, D. (2016). The crowd is a collaborative network. In P. Bjørn & J. Konstan (Eds.), *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work and Social Computing, CSCW '16,* San Francisco, USA, 27 February–2 March 2016 (pp. 134–147). New York: ACM Press.

Gupta, N., Martin, D., Hanrahan, B. V., & O'Neill, J. (2014). Turk-life in India. In D. W. McDonald & P. Bjørn (Eds.), *Proceedings of the 18th International Conference on Supporting Group Work, GROUP '14,* Sanibel Island, USA, 9–12 November 2014 (pp. 1–11). New York: ACM Press.

Haas, D., Ansel, J., Gu, L., & Marcus, A. (2015, August). Argonaut: Macrotask crowdsourcing for complex data processing. In *Proceedings of the VLDB Endow* (Vol. 8, no. 12, pp. 1642–1653). http://dx.doi.org/10.14778/2824032.2824062.

Heikkinen, T., Goncalves, J., Kostakos, V., Elhart, I., & Ojala, T. (2014). Tandem browsing toolkit: Distributed multi-display interfaces with web technologies. In A. Quigley (Ed.), *Proceedings of the International Symposium on Pervasive Displays, PerDis '14,* Copenhagen, Denmark, 3–4 June 2014 (pp. 142–147). New York: ACM Press.

Heimerl, K., Gawalt, B., Chen, K., Parikh, T., & Hartmann, B. (2012). Community sourcing: Engaging local crowds to perform expert work via physical kiosks. In H. Chi & K. Höök (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '12,* Austin, USA, 5–10 May 2012 (pp. 1539–1548). New York: ACM Press.

Hosio, S., Goncalves, J., & Kostakos, V. (2013). Application discoverability on multipurpose public displays: Popularity comes at a price. In *Proceedings of International Symposium on Pervasive Displays 2013 (PerDis'13)* (pp. 31–36).

Hosio, S., Goncalves, J., Kostakos, V., & Riekki, J. (2015a). Crowdsourcing public opinion using urban pervasive technologies: Lessons from real-life experiments in Oulu. *Policy & Internet, 7*(2), 203–222.

Hosio, S., Goncalves, J., Lehdonvirta, V., Ferreira, D., & Kostakos, V. (2014). Situated crowdsourcing using a market model. In M. Dontcheva & D. Wigdor (Eds.), *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology, UIST '14, Honolulu, USA,* 5–8 October 2014 (pp. 55–64). New York: ACM Press.

Hosio, S., Harper, R., O'Hara, K., Goncalves, J., & Kostakos, V. (2015). Life through the lens: A qualitative investigation of human behaviour with an urban photography service. In P. Olivier & D. Foster (Eds.), *Proceedings of the 2015 British HCI Conference, British HCI '15,* Lincoln, United Kingdom, 13–17 July 2015 (pp. 157–164). New York: ACM Press.
Hosio, S., Kukka, H., Goncalves, J., Kostakos, V., & Ojala, T. (2016). Toward meaningful engagement with pervasive displays. *IEEE Pervasive Computing, 15*(3), 24–31.
Huan, Y., Shema, A., & Xia, H. (2017). A proposed genome of mobile and situated crowdsourcing and its design implications for encouraging contributions. *International Journal of Human-Computer Studies, 102,* 69–80.
Huang, Y.-C. (2015). Designing a micro-volunteering platform for situated crowdsourcing. In L. Ciolfi & D. McDonald (Eds.), *Proceedings of the 19th ACM Conference Companion on Computer-Supported Cooperative Work and Social Computing, CSCW '15,* Vancouver, Canada, 14–18 March 2015 (pp. 73–76). New York: ACM Press.
Ipeirotis, P. G., & Gabrilovich, E. (2014). Quizz: Targeted crowdsourcing with a billion (potential) users. In A. Broder, K. Shim, & T. Suel (Eds.), *Proceedings of the 23rd International Conference on World Wide Web, WWW '14,* Seoul, South Korea, 7–11 April 2014 (pp. 143–154). New York: ACM Press.
Jurmu, M., Goncalves, J., Riekki, J., & Ojala, T. (2014). Exploring use and appropriation of a non-moderated community display. In S. W. Loke, L. Kulik, & E. Pitoura (Eds.), *Proceedings of the 13th International Conference on Mobile and Ubiquitous Multimedia, MUM '14,* Melbourne, Australia, 25–28 November 2014 (pp. 107–115). New York: ACM Press.
Kittur, A. (2010). Crowdsourcing, collaboration and creativity. *XRDS: Crossroads, The ACM Magazine for Students*, *17*(2), 22–26.
Kittur, A., Nickerson, J. V., Bernstein, M., Gerber, E., Shaw, A., Zimmerman, J., et al. (2013). The future of crowd work. In C. Lampe & L. Terveen (Eds.), *Proceedings of the 2013 Conference on Computer-Supported Cooperative Work and Social Computing, CSCW '13,* San Antonio, USA, 23–27 February 2013 (pp. 1301–1318). New York: ACM Press.
Kuikkaniemi, K., Jacucci, G., Turpeinen, M., Hoggan, E., & Müller, J. (2011). From space to stage: How interactive screens will change urban life. *Computer, 44*(6), 40–47.
Ludwig, T., Kotthaus, C., Reuter, C., van Dongen, S., & Pipek, V. (2017). Situated crowdsourcing during disasters: Managing the tasks of spontaneous volunteers through public displays. *International Journal of Human-Computer Studies, 102,* 103–121.
Malone, T. W., Laubacher, R., & Dellarocas, C. (2010). The collective intelligence genome. *MIT Sloan Management Review, 51*(3), 21–31.
Martin, D., Hanrahan, B. V., O'Neill, J., & Gupta, N. (2014). Being a turker. In M. R. Morris, & M. Reddy (Eds.), *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW '14,* Baltimore, USA, 15–19 February 2013 (pp. 224–235). New York: ACM Press.
Martin, D., O'Neill, J., Gupta, N., & Hanrahan, B. V. (2016). Turking in a global labour market. *Computer Supported Cooperative Work, 25*(1), 39–77.
Müller, J., Alt, F., Michelis, D., & Schmidt, A. (2010). Requirements and design space for interactive public displays. In Smeulders, A. (Ed.), *Proceedings of the 18th ACM International Conference on Multimedia, MM '10,* Firenze, Italy, 25–29 October 2010 (pp. 1285–1294). New York: ACM Press.
Müller, J., Wilmsmann, D., Exeler, J., Buzeck, M., Schmidt, A., Jay, T., et al. (2009). Display blindness: The effect of expectations on attention towards digital signage. In H. Tokuda, M. Beigl, A. Friday, A. J. Brush, & Y. Tobe (Eds.), *Proceedings of the International Conference on Pervasive Computing, Pervasive '09,* Nara, Japan, 17–20 May 2010 (pp. 1–8). Berlin, Heidelberg: Springer.
Peltonen, P., Kurvinen, E., Salovaara, A., Jacucci, G., Ilmonen, T., Evans, J., et al. (2008). It's mine, don't touch!: Interactions at a large multi-touch display. In D. Tan (Ed.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '08,* Florence, Italy, 5–10 April 2008 (pp. 1285–1294). New York: ACM Press.

Retelny, D., Robaszkiewicz, S., To, A., Lasecki, W. S., Patel, J.. Rahmati, N., et al. (2014). Expert crowdsourcing with flash teams. In M. Dontcheva & D. Wigdor (Eds.), *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology, UIST '14*, Honolulu, USA, 5–8 October 2014 (pp. 75–85). New York: ACM Press.

Salehi, N., McCabe, A., Valentine, M., & Bernstein, M. (2017). Huddler: Convening stable and familiar crowd teams despite unpredictable availability. In L. Barkhuus, M. Borges, & W. Kellogg (Eds.), *Proceedings of the 2017 ACM Conference on Computer-Supported Cooperative Work and Social Computing, CSCW '17*, San Francisco, USA, 25 February–1 March 2017 (pp. 1700–1713). New York: ACM Press.

Saxton, G. D., Onook, O., & Kishore, R. (2013). Rules of crowdsourcing: Models, issues, and systems of control. *Information Systems Management, 30*(1), 2–20.

Schmitz, H., & Lykourentzou, I. (2018, January). Online sequencing of non-decomposable macro-tasks in expert crowdsourcing. *ACM Transactions on Social Computing, 1*, 1, Article 1, 33 pp. https://doi.org/10.1145/3140459.

Shen, C., Everitt, K., & Ryall, K. (2003). UbiTable: Impromptu face-to-face collaboration on horizontal interactive surfaces. In A. K. Dey, A. Schimdt, & J. F. McCarthy (Eds.), *Proceedings of the International Conference on Ubiquitous Computing, UbiComp '03,* Seattle, USA, 12–15 October 2003 (pp. 281–288). Berlin, Heidelberg: Springer.

Shirky, C. (2010). *Cognitive surplus: How technology makes consumers into collaborators.* Penguin.

Terrenghi, L., Quigley, A., & Dix, A. (2009). A taxonomy for and analysis of multi-person-display ecosystems. *Personal and Ubiquitous Computing, 13*(8), 583–598.

Weißker, T., Berst, A., Hartmann, J., & Echtler, F. (2016). The massive mobile multiuser framework: Enabling ad-hoc realtime interaction on public displays with mobile devices. In J. Muller & N. Memarovic (Eds.), *Proceedings of the 5th ACM International Symposium on Pervasive Displays, PerDis '16*, Oulu, Finland, 20–22 June 2016 (pp. 168–174). New York: ACM Press.