Please Stay vs Let's Play: Social Pressure Incentives in Paid Collaborative Crowdsourcing

Oluwaseyi Feyisetan(✉) and Elena Simperl

University of Southampton, Southampton, UK
{oof1v13,e.simperl}@soton.ac.uk

Abstract. Crowdsourcing via paid microtasks has traditionally been approached as an individual activity with units of work created and completed independently. Other forms of crowdsourcing have however, embraced a mixed model that further allows for interaction and collaboration. In this paper, we expand the model of collaborative crowdsourcing to explore the role of social pressure and social flow generated by partners, as sources of incentives for improved output. We designed experiments wherein a worker could request their partner to collaboratively complete more tasks than required, either not to be abandoned and lose money (social pressure), or for fun (social flow). Our experiments reveal that these socially motivated incentives can act as furtherance mechanisms improving output by over 30 % and accuracy by about 5 %.

Keywords: Paid microtask crowdsourcing · Collaboration · Social pressure · Social flow · Incentives engineering

1 Introduction

Microtask crowdsourcing is one of the most prominent forms of online crowdsourcing [5]. It is primarily used when the work to be outsourced is highly parallelizable and can be divided into smaller pieces down to a micro level, which takes only seconds to minutes to complete. It brings together two sets of actors: requesters (the persons or institutions seeking help from the crowd) and workers (individuals or teams taking on tasks advertised via an open call).

In its most typical instance, microtask crowdsourcing is understood as an aggregation of individual contributions that have been created independently of each other [5]. While this model has some advantages, allowing large numbers of people to take on tasks and complete significant amounts of them quickly,[1] there are also situations in which allowing for social interaction or collaboration is preferred [1,11]. This is a function of many factors: from the complexity of the work to be outsourced [12]; to using synchronicity of crowd answers as a means to

[1] For instance, the citizen science project Subspotter achieved 350000 classifications in 20 h with the help of their community of volunteers, see http://blog.zooniverse.org/2015/08/31/. For paid microtasks requesters expect answers in their thousands or more within hours or days.

A. Bozzon et al. (Eds.): ICWE 2016, LNCS 9671, pp. 405–412, 2016.
DOI: 10.1007/978-3-319-38791-8_26

drive participation [16]; or giving the crowd a place to self-organize, share ideas and experiences, and discuss [15]. In this paper, we present results of experiments on paid crowdsourcing, carried out on a bespoke platform - *Wordsmith*, featuring elements of interaction and collaboration. We study the effect of social pressure and social flow on the task quality and output of workers carrying out tasks in this setting.

2 Background and Related Work

Some GWAPs (Games with a purpose), for example, von Ahn's ESP game [16] have a strong element of interaction among contributors. These games employ various strategies such as output agreement between players to generate useful work results and drive engagement. However, in microtask models that involve financial payment, which do not employ these strategies, it becomes important to understand and leverage on these factors (such as collaboration) that might drive participation.

2.1 Factors Affecting Collaborative Work

Several works have looked at the effect of peer and social pressure in incentivising work output collaboratively online and in business enterprises. The effects of peer pressure on contributions to enterprise social media were studied by [2] where they observed that, the participation of a worker's manager is a key source of social pressure in initiating contribution while, comments fuel the pressure for sustained contributions. From their work, we observe that most of the social pressure effects that are seen in the offline world are present, and amplified in the online world.

Social flow stemmed out of an extension of Csikszentmihalyi's 'theory of optimal experience' [4] where flow (or individual/solitary flow), was presented as an intrinsically rewarding, highly absorbing state, which is attainable when individuals freely choose an activity with: clear goals, immediate feedback, and a balance between challenge and skills. Despite the freedom and pleasure that comes from immersive individualistic activities, it has been observed that some of the most gratifying flow experiences occur in social circumstances [7], leading to the concept of *social flow*. Walker [17] presented the conditions and indicators of *social flow*. Some conditions stated include: immediate and clear feedback from the task and group members, interdependence and cooperation, and the challenges being important to the whole group.

2.2 Our Contribution

We look at the impact of socially motivated incentives of pressure and flow as drivers for improved quality and task output in paid collaborative microtask crowdsourcing. We show that beyond basic collaborative crowdsourcing, we can improve task and performance results by employing empathy-centric social pressure and the desire for social fun.

3 The Wordsmith Platform

The Wordsmith microtask crowdsourcing platform consists of an image labelling task to be carried out either by a single worker, or collaboratively by two workers. The collaborative mode in Wordsmith is reminiscent of the ESP game [16] which was designed for voluntary unpaid players who were not primarily motivated by money. The heart of Wordsmith consists of images displayed either sequentially to a single worker, or displayed in-sync to two workers. The workers are in turn, required to input multiple informative keywords (also referred to as labels or tags), which describe the image into a text field. Each keyword is accepted as *valid* based on a simple computed measure of quality: the keyword must be in English (validated via a dictionary web service) and it must not have been used by the worker to describe either that image or repeatedly for previous images - else, all the images could be labeled as *cat, cat, cat.*

We incorporated social and non-social feedback elements to Wordsmith to provide workers with real-time and longitudinal information on their progress in the task. Providing feedback has been shown to improve retention and engagement by enhancing intrinsic feelings of accomplishment as people advance in a task [4,14]. The platform included visual feedback on the number of images annotated, the number of tags generated and how each worker's output compares with that of others. More details on Wordsmith's gamification features and design philosophy can be found in our earlier work here [6].

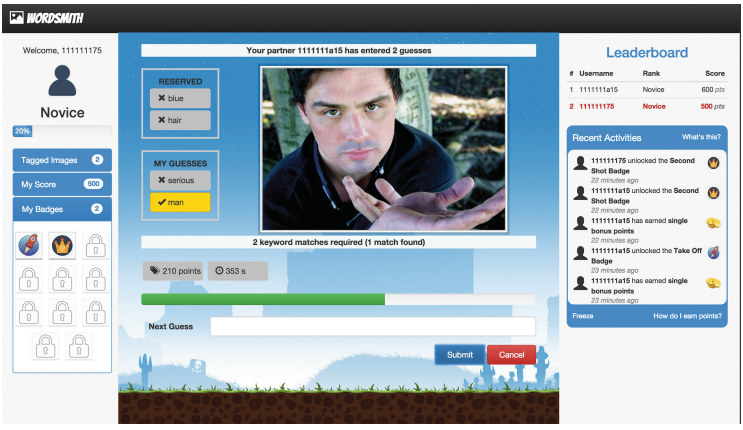

Fig. 1. Wordsmith interface

4 Social Pressure in Wordsmith

When the *social pressure* setting is activated in the collaborative mode in Wordsmith, a worker is given a heads up when their partner is about to quit the task.

The worker can then select one of two options: *tell them to stay* or *allow them to go*. Choosing to request their partner to remain in the task represents the cost of exerting social pressure as presented by [3]. The message which is then automatically conveyed to the partner is dependent on the sending worker's current level in the task: the message either appeals to the partner to continue till the requesting worker reaches the level where they receive a payoff (*please stay*) or, the message requests that both workers continue annotating for fun if the payoff has been received by the requesting worker (*let's play*).

5 Experimental Design

We carried out a study in which a number of workers were recruited from a large pool, and required to annotate images either in the traditional single worker mode (**SP**) or the collaborative mode (**MP**) of Wordsmith. Within each task mode, workers were required to annotate a certain number of images in order to get paid: annotate 1 image (**LT** - *low threshold*) or annotate 11 images (**HT** - *high threshold*) based on earlier work [6]. Finally, in the collaborative, high-threshold mode (**MP-HT**), we carried out sub studies as follows: in one condition, a partner attempting to exit the task was allowed to leave, in another condition, a partner attempting to exit the task could be shown a **please stay**, or **let's play** message at the request of their partner.

5.1 Research Questions

1. Does collaboration work as an effective model for paid microtask?
2. What is the role of social pressure incentives in collaborative tasks?

6 Results

We recruited participants from CrowdFlower[2]: 600 workers for the single worker experiments modes and 600 workers for the collaborative experiments.

6.1 Social Pressure: Please Stay vs Let's Play

When Wordsmith senses that a worker's partner is about to exit the task, the worker is alerted. The worker can then request their partner to remain in the task. If the worker has tagged less than the requisite number of images in order to get paid, the worker can send a *please stay* request, else, the worker can send a *let's play* request. The receiving worker can then decide to stay (*i will stay*) or to leave the task (*i will go*) Table 1.

Social Pressure Requests. From the results in Fig. 2 with its accompanying table, we observe that workers are more likely to initiate a request of any kind

[2] https://crowdflower.com.

Table 1. *Experiment results* - summary of experiment results

Experiment results					
	Low threshold		High threshold		
	Traditional	Collaborative	Traditional	Collaborative	Soc'l pressure
Total workers	402	365	514	499	508
Total tags	21,538	48,171	27,652	108,950	158,716
Unique images tagged	200	200	2,196	2,200	2,200
Inter-annotator agreement	29.44 %	**34.55**%	14.26 %	25.82 %	**29.35**%
ESP tags agreement	41.26 %	25.39 %	43.96 %	37.94 %	40.11 %
Avg images tagged/person	26.68 (SD = 38.21)	9.77 (SD = 13.23)	26.75 (SD = 42.07)	25.05 (SD = 17.92)	29.00 (SD = 28.30)
Avg tags/person	53.57	**131.97**	53.80	218.34	**312.43**
Avg new tags/person	2.78 (1,117/402)	**8.69** (3,172/365)	1.80 (925/514)	11.83 (5,903/499)	**16.21** (8,236/508)

when they have not been paid. When a worker has not yet been paid, they are more likely to request that their partner stay (*please stay request*) than permitting their partner to leave. After the workers had been paid, they were also more likely to request their partner to stay (*let's play request*) than permitting them to leave. The *please stay* requests ($Requests = 1,023$) were used more frequently as a social incentive than the *let's play* request ($Requests = 151$), suggesting that workers are more inclined to put pressure on their partners when there is financial reward at stake than just fun. Figure 2 also reveals that some workers would actually release their partner to leave and wait to be connected to another partner. It shows that on the average, as expected, fewer workers (20 % *vs* 35 %) who haven't been paid would opt for this option.

Social Pressure Responses Figure 3 summarises the results (in a logarithmic scale) of a worker's responses to both *please stay* and *let's play* requests. When a worker receives a *please stay* request (signifying that the requesting partner has not yet been paid), they can respond by choosing either to stay (*I will stay*) or to leave (*I will go*). The choice to stay or to leave also varies depending on whether

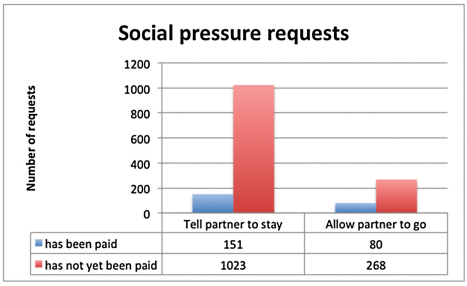

Fig. 2. Social pressure requests made by workers before and after payment (Color figure online)

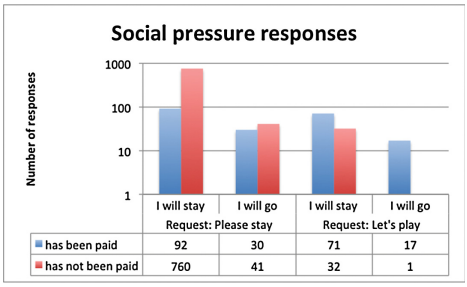

Fig. 3. Worker responses to *please stay* and *let's play* requests (on log scale) (Color figure online)

the receiving worker has been paid or not. The results indicate that, a worker who has not been paid, receiving a *please stay* message from a fellow unpaid is more likely to stay, with 95 % probability, than to exit the task (760 *vs* 41). This is in line with workers being incentivised by having shared circumstances (i.e., the need to both get paid), as stated by [9]. Similarly, a worker receiving a *please stay* request from an unpaid worker, after they have been paid, is also likely to respond by staying, albeit, with a slightly less probability of 75 % (92 *vs* 30). Furthermore, a worker receiving a *let's play* request (from a worker that has been paid) can also choose to stay or to leave, depending on whether the receiving worker has been paid or not. The results illustrate that, a worker who has not been paid, previously intending to exit the task, would almost certainly remain in the task after being sent a *let's play* message with 97 % probability (32 *vs* 1). The result also reveals the response to social flow incentives: a worker who has been paid would return to continue playing with another worker with 80 % certainty, even more likely than they would help a partner get paid (although, the results suggest that these requests occur less frequently). This is also another form of incentivisation by having shared circumstances (i.e., the desire to re-experience social flow).

Figure 3 also gives insights into when workers decide to leave their partners, despite receiving either a *please stay* or *let's play* request. The results reveal that, after receiving a *please stay* request from a worker who has not been paid, a receiving worker is more likely to leave if they have not been paid also. Hence they do not feel any guilt from leaving their partner hanging since they haven't been paid also. Similarly, after receiving a *let's play* request from a partner who has been paid, the receiving worker is more likely to decline the offer and choose to exit the task if they have also been paid.

7 Discussion

Collaboration Must Pay Off. The benefits of collaborative participation in the image labelling task were more visible in the *high threshold* conditions. In the traditional mode, workers annotated, on average, the same number of images

(and generated the same number of tags) in the high and low threshold conditions. In other words, without the restriction of partner agreements, the task threshold did not really make a difference. In the low task threshold condition of the collaborative mode, workers tagged more than the requisite number of images, nevertheless, this positive delta was not sufficient to match up to the individual freedom afforded in the traditional mode. The high task threshold on the other hand indicates in the collaborative setting, how the power of (and aspiration towards) social concordance, propped up by a higher payment cutoff can be leveraged to generate more and better results. Workers in this condition, initially motivated by the need to get paid, worked together to realise improved results. This finding contributes to the larger discussion around motivation and paid microtask crowdsourcing. Surveys such as [6,10,13] have observed that financial incentives are just one, though important, part of a much more refined story of motivation of workers. The present work offers evidence on the effects of social pressure and flow, but it also raises new questions regarding the implications of the findings for incentives design that take into account particular types of workers (e.g., top contributors vs casual visitors).

Workers Behave Empathically. Social pressure incentives could be harnessed to attain speedy task completion and encourage empathic collaboration. Our analysis revealed that workers on realising that their partners have not reached the task threshold for payment, would be willing to annotate a few more images to help them get paid. This is in contrast to the individualistic thinking model which has been enshrined in traditional paid micro task platform settings. Our results show that paid workers would be willing, not only to work together, but to go the extra mile to ensure that their partner also gets paid. Workers respond not only to the need to help their partner get paid, they also respond to their partner's desire to continue annotating just for the fun of it. As noted earlier, while these results are encouraging, to develop a theory of incentives for paid microtask platforms, one would need additional experiments that take into account worker behavior patterns, as well as other tasks and possibly more complex collaboration models.

Ethics and Compensation. We understand the negative effects extended gameplay can have on people [8], especially on those who might depend on crowdsourcing as a supplementary income source. We attempted to keep our payments at the fair amount of $0.10 per minute and our low threshold condition paid up to $7.2 per hour ($0.02 for 1 image).

8 Conclusion

This paper presents results showing that social incentives could be used to boost the performance of participants in collaborative crowdsourcing. These results are in line with findings from GWAPs and multi-actor crowdsourcing systems, and could be used to inform the re-design of paid microtask platforms such as CrowdFlower and Mechanical Turk which do not integrate collaborative workflows as first-class citizens.

References

1. Bernstein, M.S., Brandt, J., Miller, R.C., Karger, D.R.: Crowds in two seconds: enabling realtime crowd-powered interfaces. In: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology
2. Brzozowski, M.J., Sandholm, T., Hogg, T.: Effects of feedback and peer pressure on contributions to enterprise social media. In: Proceedings of the ACM International Conference on Supporting Group Work (2009)
3. Calvó-Armengol, A., Jackson, M.O.: Peer pressure. J. Eur. Econ. Assoc. **8**(1), 62–89 (2010)
4. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience, vol. 41. Harper Perennial, New York (1991)
5. Dawson, R., Bynghall, S.: Getting Results from Crowds. Advanced Human Technologies, San Francisco (2012)
6. Feyisetan, O., Simperl, E., Van Kleek, M., Shadbolt, N.: Improving paid microtasks through gamification and adaptive furtherance incentives. In: Proceedings of the 24th International Conference on World Wide Web (2015)
7. Jackson, S.A., Csikszentmihalyi, M.: Flow in Sports. Human Kinetics, Champaign (1999)
8. Joukhador, J., Blaszczynski, A., Maccallum, F.: Superstitious beliefs in gambling among problem and non-problem gamblers: preliminary data. J. Gambl. Stud. **20**(2), 171–180 (2004)
9. Kandel, E., Lazear, E.P.: Peer pressure and partnerships. J. Polit. Econ. **100**, 801–817 (1992)
10. Kaufmann, N., Schulze, T., Veit, D.: More than fun and money. worker motivation in crowdsourcing-a study on mechanical turk. In: AMCIS, vol. 11, pp. 1–11 (2011)
11. Kittur, A.: Crowdsourcing, collaboration and creativity. ACM Crossroads **17**(2), 22–26 (2010)
12. Kittur, A., Smus, B., Khamkar, S., Kraut, R.E.: Crowdforge: crowdsourcing complex work. In: Proceedings of the 24th Annual ACM symposium on User Interface Software and Technology, pp. 43–52. ACM (2011)
13. Mason, W., Watts, D.J.: Financial incentives and the performance of crowds. ACM SIGKDD Explor. Newsl. **11**(2), 100–108 (2010)
14. McGonigal, J.: Reality is Broken: Why Games Make Us Better and How They Can Change the World. Penguin Group, London (2011)
15. Tinati, R., Van Kleek, M., Simperl, E., Luczak-Rösch, M., Simpson, R., Shadbolt, N.: Designing for citizen data analysis: a cross-sectional case study of a multi-domain citizen science platform. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (2015)
16. von Ahn, L.,Dabbish, L.: Labeling images with a computer game. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2004, pp. 319–326. ACM, New York, NY, USA (2004)
17. Walker, C.J.: Experiencing flow: is doing it together better than doing it alone? J. Positive Psychol. **5**(1), 3–11 (2010)