

All Onboard: The Effects of Voice Communication on Online Collaboration

Matthew Pick, Tope Olagbemi, Kartik Sawhney

Stanford University

{mpick, topeo, kartiks2}@stanford.edu

ABSTRACT

We were inspired by the previous work of the Flash Orgs Research Team at Stanford University with the Foundry system for crowd workers [9]. Although Foundry has been used to coordinate individual, asynchronous tasks, it has never been used for synchronous collaboration. To better understand how to improve the output of these crowd teams, we presented teams of three workers with the task of creating an online advertisement for an audio recording car radio where each worker was assigned a specific role: Artist 1, Artist 2, or Text Creator. The experimental teams were led through a brief voice chat where each team member introduced himself/herself and described his/her background, while the control groups immediately began the task. After the creation of the advertisements, we had hundreds of people vote on the advertisement images they thought were most effective at conveying the product idea. We found that the experimental advertisements received a total of 44% more votes than those in the control group. We also saw greater communication and collaboration in the experimental teams. Thus, pre-task communication not only improved the performance, but also team dynamics.

INTRODUCTION

Successful online collaboration is challenging. Regardless of what one works on, a lot of the same issues arise. Geographical and language barriers between workers from different parts of the world make effective communication difficult. Differences in skill, background, and personal beliefs can prevent teams from accomplishing all that they want to accomplish, or create friction among team members; in the case of the online encyclopedia Wikipedia, "edit wars" are regularly known to ensue between users with differing views over how a particular page should be edited [2]. Such conflicts can make effective communication even more challenging and prevent online teams from completing their task with the highest-possible quality.

To alleviate these issues, we considered ways we could make the online collaboration process less virtual and more focused on forming positive human connections. Initially, we wanted to explore the effects of voice communication on online collaboration as it pertained to skilled online workers. To do this, we hired 12 skilled freelance workers from Upwork (8 artists, 4 writers) and assigned them the task of using Foundry to help them create an advertisement for a car radio with recording and playback capabilities. At the scheduled project time, however, only 2 of the 12 hired

workers were online; as a result, we could not run the study, and we encountered major logistical issues that made it extremely difficult for us to use Upwork freelancers in the future. Thus, we had to pivot our study and have our experiment focus not on skilled online workers, but on online workers in general.

In our study, we explored the effects of having online collaborators partake in a preliminary onboarding team orientation -- during which they got to introduce themselves and briefly discuss their backgrounds -- and determined whether or not this initial orientation made a difference in the quality of their final product. All teams in our study were still tasked with using Foundry as an aide in the creation of an advertisement for a car radio with recording and playback capabilities. Our purpose in conducting this study was to provide more insight into possible ways that the process of online collaboration can be made more comfortable for those involved.

The rest of this paper discusses the specific aspects of our study. We begin by discussing the related work in the field, followed by a discussion of our hypothesis for our experiment. We then discuss our methodology for carrying out our project. After that, we analyze our findings from both a quantitative and qualitative standpoint, discussing what we would have done differently given more time. Finally, we discuss the implications that our study has both on Foundry as a platform and on online collaborative crowd work in general.

RELATED WORK

Given our focus on the impact of communication on team dynamics and final quality of work in virtual environments, our project relies heavily on crowdsourcing as well as management literature. Even though we do not directly work with crowd workers in this project, our participants are representative of unskilled crowd workers on various portals such as Amazon Mechanical Turk, making crowdsourcing literature relevant to our project.

Crowdsourcing Literature

There has been an increasing trend to hire crowd workers for long creative tasks, unlike the small and specific tasks that they have usually been hired for in the past. The Polymath project [3] and the Pipeline project [7], are two of several examples where workers have worked on complex mathematical and management-oriented tasks without any predefined workflows. However, most of these projects are

domain-specific -- mathematics and management in this case. In 2014, Retelny et al. at Stanford University developed Foundry, an interface that supports crowd workers to collaborate and work on complex interdependent versatile tasks.

To evaluate the effectiveness of Foundry, the "Flash Teams" research team brought together a group of crowd workers to design napkin sketches without any predefined task flows or instructions. While it was evident that the teams using Foundry were more efficient, there was a lot of fluctuation in the performance of the teams in general, leading to unpredictability. Nevertheless, the team claimed to observe a correlation between team dynamics and performance. We build on this observation in this paper, attempting to observe factors that lead to better team dynamics and ultimately performance.

Management Literature

Onboarding has been identified to be crucial for team dynamics and final quality of work, and is well-accepted in the industry and academia. Mikawa et al. observed it to be even more important for teams with diverse backgrounds, which is often the case with crowd workers [8]. Moreover, Dunn et al. observed promoting communication as the most important component of the onboarding process for virtual teams. Given these observations, we analyze the role of communication in team dynamics and quality of work, with particular reference to online workers.

Hypothesis on the Effects of Voice Communication

For our study, we hypothesize that a project team when given the opportunity to partake in an onboarding team orientation -- during which they use voice chat to introduce themselves to their teammates -- will improve communication among team members, give team members a more positive impression of the task, and improve the quality of that team's final product.

Our motivation for our hypothesis comes from prior findings discussing the relationship between team onboarding and team dynamics. A study conducted by Ganzel (as cited in [5]) showed that a positive correlation likely exists between the process of onboarding and a team's productivity; other studies have shown that an onboarding process centered on an individual's identity had positive effects on team morale and performance [1]. There have also been studies done with findings showing that voice chat is an authentic way to communicate during a collective activity in which the participants are remote (Geerts, 2006). However, studies have yet to be conducted regarding effective onboarding procedures for synchronous crowd workers.

To test our hypothesis, we divided teams into two separate groups. Teams in the first group (our control group) were simply told to begin working on the project and to use Foundry's text chat feature to communicate; teams in the second group (our experimental group) went through the onboarding voice chat orientation, then switched over to

Foundry's text chat tool to communicate and begin working on the project. To ensure that teams accurately reflected teams in an online collaborative environment, we constructed the teams in such a way that all the members on a team were complete strangers to one another and the roles as Artist 1, Artist 2, and Text Creator were randomly distributed prior to the task beginning. Once the maximum amount of time that a team could spend on the project had passed, we conducted brief exit interviews with each of the team members, asking them about their experience working with their team on the project. We also qualitatively analyzed the communication patterns in Foundry's chat logs for each of our teams.

METHODOLOGY

To observe the role of communication, we conducted two pilot studies with a total of six teams of three people each. We avoided large teams due to the logistical challenges involved, yet we wanted the teams to be large enough and the roles interwoven such that team members were encouraged to communicate. Unfortunately, we faced several challenges in getting crowd workers to participate in our study, and so recruited Stanford students, other college students located across the nation, and international participants in India. Since diversity is a key characteristic of crowd work, we ensured that our teams were very diverse. As an example, one of our teams consisted of an Indian college student, an American college student and a middle-aged American.

These six teams were divided into two categories: the control group and the experimental group. Each of these teams was asked to design an advertisement for a car radio with recording and playback capabilities in 45 minutes. A team consisted of two artists (Artist 1 and Artist 2) and a Text Creator, and all of them worked together to develop a single online advertisement. The teams used Foundry (the platform developed by the "Flash Teams" research team at Stanford) to communicate and share their work with other members. Its support for collaboration made Foundry ideal for our workers.

In addition to Foundry as a platform, the teams in the experimental group also had an opportunity to participate in a brief telephonic call before the task. This was a platform for workers to introduce themselves and talk about their background and expectations. These calls were closely monitored to ensure that the workers do not discuss anything related to the actual task that could give them an edge over the teams in the control group.

Qualitative Observations

During the study, we noticed interesting differences between teams. Two out of the three control groups immediately began working on a design with very little interaction between team members. They chose not to introduce themselves and instead performed their assigned tasks,

occasionally reaching out to the group for feedback on their design choices. While the third group solicited feedback more often, the members did not introduce themselves either. The chat logs revealed relatively little collaboration.

On the other hand, the experimental group had very different observable experiences. Following the initial phone call, the teams spent a significant amount of time brainstorming and planning the project. While each team member took ownership of his individual task, there was far more collaboration with other members. In one instance, the team members decided to switch roles based on their background and skills discussed during the phone call.

To get feedback about their experiences, we also conducted brief exit interviews for all participants. Most participants were satisfied with the system (Foundry), while some reported bugs in the system; these bugs have been reported to the Foundry team. As regards their experiences working with their team, the experimental group felt more comfortable working with the people. All the participants on this group appreciated the phone call, and in fact, suggested to transition to a video call. As one of the participants put it, "it [the phone call] added personality to the participants". Due to more collaboration, members on these teams also seemed to have higher expectations from their co-workers, and so sometimes were critical of each other. All of the participants claimed that they would have introduced themselves even if there was no phone call, which is interesting since none of our control groups did so.

Teams in the control group often had complaints about miscommunication, but had little to say about the other team members. This seems to be because of lack of adequate communication and collaboration. Moreover, one of the members on two of the three teams in this group stopped communicating after a while, later expressing embarrassment at what they believed to be their own incompetence. Most members thought that a phone call would have been very helpful, though some thought that the task wasn't complex enough to warrant it. Given these differences between the experimental and control groups, it seems that the phone call had a positive effect on the team dynamics and promoted collaboration.

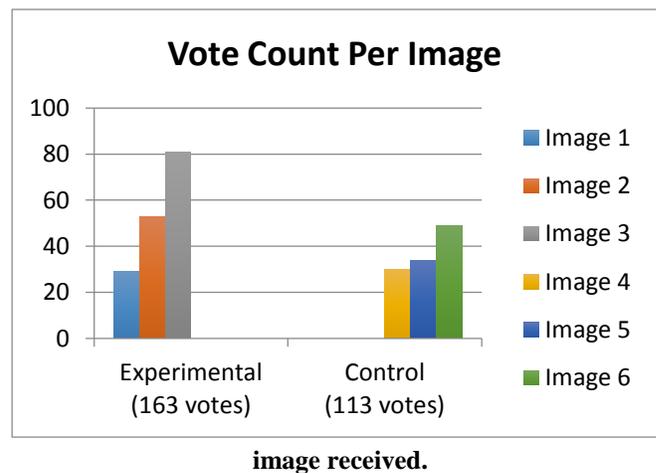
To better understand the impact of communication on the final quality of the image produced, the images were analyzed by a Business Manager in the Marketing Solutions advertising department at LinkedIn for professional feedback. She was not told which of the two sets of images were experiential and control so as to not influence her analysis. Her feedback was as follows: "[The experimental group's] images are better quality in terms of design and brand recognition. Simplicity, sleekness and better image resolution helps consumers easily recognize it, and it seems more legitimate to last over time....In the long term when establishing a brand, [the experimental group's] design would resonate and be recognized more over time, which would create more longevity for the company (i.e. the Nike

swoosh)." It is evident then from these observations that there is a positive correlation between effective communication and team performance in online teams.

Quantitative Analysis

In order to quantitatively analyze the advertisements, we created our own custom website and solicited votes from students at Stanford and through Mechanical Turk. The site randomly displays, side-by-side, one image from the control group and one from the experimental group, randomizing the order of the two images. Voters then click on the image they feel is the better advertisement for the audio recording car radio. Our voting site also contains an admin interface that shows the total number of votes for the experimental and control groups, as well as the breakdown of the votes per image. Figure 1 shows the total votes per image in each of the two groups. The control images received a total of 113 votes and the experimental ones received 163 votes.

Figure 1: Chart detailing the number of votes each



We conducted a Chi Squared test with the Control and Experimental Groups as our categories, 113 observed in the Control and 163 observed in the Experimental, and 138 expected votes for each category. This results in a Chi squared value of 9.058 with 1 degree of freedom and a P value of 0.0026. This means that our results are statistically significant with a 5% significance threshold. Since there was some variation among image quality within the two groups, we also conducted an unpaired t-test over the individual images within the categories of experimental and control, which resulted in a two-tailed P value of .52. This P value is to be expected due to the small n value of 3 per condition. In order to even possibly achieve statistical significance with the t-test, we would need to have a much larger n value.

Given more time, we would run more studies and, thus, significantly increase our sample size. We would also solicit more votes so that image quality would play less of a factor in determining which images get more

votes.

Implications of Our Results

Although our t-test did not yield statistical significance, the Chi Squared test, qualitative results from the participant interviews, and LinkedIn professional analysis indicate that it may be beneficial to hold a brief voice conversation among crowd workers. Such information is useful in guiding how platforms such as Foundry should be built in the future to encourage collaboration among teammates that have no prior relationship. This voice conversation could be built into the system as part of an orientation before the task begins where workers introduce themselves and begin discussing the tasks. Although several of our study participants who were all US college students said they would have preferred a video call, the Upwork workers we recruited felt uncomfortable video conferencing and instead preferred the audio only medium. Therefore we do not believe Foundry should incorporate video calls as it almost entirely recruits Upwork workers for tasks.

While our results are regarding synchronous crowd work, given our experience conducting this study, it is doubtful whether this is a field worth pursuing for flash teams. As previously mentioned, it was extremely difficult to organize workers to begin at the same time as the vast majority either didn't show up or began the task late. This would prevent any larger scale tasks that require workers to synchronously collaborate or interact via a voice call. Even though our results show the possible advantages of real time voice communication, the unexpected and ironic results to come from our study may be that it is implausible to proceed with this line of synchronous crowd work research.

CONCLUSION

Drawing inspiration from crowdsourcing and management literature, this paper analyzed the role of communication on team dynamics and performance in online teams. We conducted pilot studies with six teams-three of the teams simply used text chat to communicate, while the other three got an opportunity to participate in a short telephonic conversation. It is evident from the qualitative and quantitative results that the short phone conversation (which was restricted to introductions, background and strengths and weaknesses etc.) had a positive impact on not only the team dynamics but also the final performance of the team (as determined from the quality of the images). While we realize the sample size to be small to make a claim, the paper offers interesting results that can guide effective on-boarding practices for crowd work.

ACKNOWLEDGMENTS

We would like to thank all of the individuals who volunteered to participate in our study. We would also like to thank the individuals who provided regular guidance and feedback for us as we conducted this research.

ADVERTISEMENTS CREATED

Experimental Group Images:

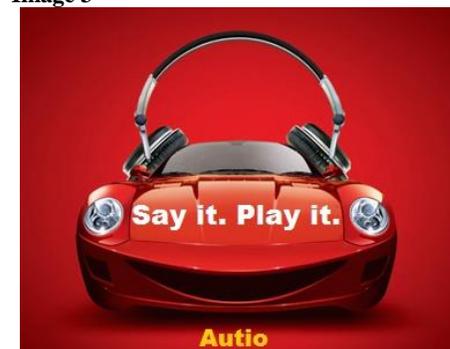
Image 1



Image 2



Image 3



Control Images

Image 4



Image 5



Image 6



REFERENCES

1. Cable, D. M., Gino, F., & Staats, B. R. (2012). Reinventing Employee Onboarding. *Image*.
2. Chen, J., Ren, Y., & Riedl, J. (2010, April). The effects of diversity on group productivity and member withdrawal in online volunteer groups. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 821-830). ACM

3. Cranshaw, J., & Kittur, A. (2011, May). The polymath project: lessons from a successful online collaboration in mathematics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1865-1874). ACM.
4. Dunn, S., Grannan, C., Raisinghani, M., & Stalling, H. Communication Strategies for Successful Virtual Teams.
5. Dai, G., & De Meuse, K. P. (2007). A review of onboarding literature. *Lominger Limited, Inc., a subsidiary of Korn/Ferry International*.
6. Geerts, D. (2006, October). Comparing voice chat and text chat in a communication tool for interactive television. In *Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles* (pp. 461-464). ACM.
7. Luther, K., Fiesler, C., & Bruckman, A. (2013, February). Redistributing leadership in online creative collaboration. In *Proceedings of the 2013 conference on Computer supported cooperative work* (pp. 1007-1022). ACM.
8. Mikawa, S. P., Cunnington, S. K., & Gaskins, S. A. (2009, February). Removing barriers to trust in distributed teams: understanding cultural differences and strengthening social ties. In *Proceedings of the 2009 international workshop on Intercultural collaboration* (pp. 273-276). ACM.
9. Retelny, D., Robaszekiewicz, S., To, A., Lasecki, W. S., Patel, J., Rahmati, N., ... & Bernstein, M. S. (2014, October). Expert crowdsourcing with flash teams. In *Proceedings of the 27th annual ACM symposium on User interface software and technology* (pp. 75-85). ACM.