Building Empathy: Scaling User Research for Organizational Impact

Ariel Liu

Google, Inc.
Mountain View, CA 94043, USA
arielliu@google.com

Victoria Schwanda Sosik

Google, Inc.
New York, NY 10011, USA
victoriasosik@google.com

Khadine Singh

Google, Inc.
Mountain View, CA 94043, USA
khadine@google.com

Abstract

Building user empathy in a tech organization is crucial to ensure that products are designed with an eye toward user needs and experiences. The Pokerface program is a Google internal user empathy campaign with 26 researchers that helped more than 1500 employees—including engineers, product managers, designers, analysts, and program managers across more than 15 sites—have first-hand experiences with their users. Here, we discuss the goals of the Pokerface program, some challenges that we have faced during execution, and the impact we have measured thus far.

Author Keywords

Empathy building; usability study; industry research.

ACM Classification Keywords

H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

Introduction and Goals

Research has shown that user involvement throughout the entire innovation process has been a key to success in many high-tech firms [3]. Feedback from users has helped companies uncover key user needs as well as understand how users accomplish critical journeys with products. However, any given research presentation is limited in how well

---

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).

CHI'18 Extended Abstracts, April 21–26, 2018, Montréal, QC, Canada.

ACM ISBN 978-1-4503-5621-3/18/04.

<http://dx.doi.org/10.1145/3170427.3174352>

it can convey users' issues to the teams building products. To fully engage with and understand the spectrum of users' problems and needs, all employees involved in product development must get to know their users. Building user empathy is critical—it is essential for developers to be exposed to others' perspectives in order to better understand them [5].

To date, several programs have been created to allow product teams to better understand their users, especially around how teams can integrate user emotions and values into their product development process [1] [4]. Similar to these established programs, our objective is to increase user awareness with our teams and acknowledge the fact that "we aren't the users", but by going a step further and putting product teams face to face with their users. Towards this goal, we created the Pokerface program.

Pokerface is a Google internal user empathy program. It provides a seminar on the main principles of user experience research as well as hands on experience moderating a user research study session with users and a debrief session with the team to discuss their findings. The primary goal is to build user empathy across Google's engineers, product managers, designers, and program managers by having the teams conduct their own user research sessions, interacting with users face-to-face. Our secondary goal is twofold: exposing other job functions to the research process and our team of researchers and gathering impactful insights that teams can immediately act on. By doing so we strive not to just provide teams with one dose of empathy, but to show them the value of user research and to inspire them to work with UX researchers throughout their product life cycles.

This program is based on an earlier Google project that ran from 2012–13, which saw many successes, includ-

ing general excitement around user-centered research, breaking preconceived perceptions about products, and bug fixing [2]. While this program was successful, it was somewhat limited in scope. It was successful in reaching over 400 Googlers from more than 6 sites, but this time—4 years later—our goals were much larger, targeting more than 1500 Googlers in 15 offices across the world. The increased complexity of our project brought along many additional challenges which we will discuss in this case study. In this iteration of the program we also executed a more thorough evaluation of the program's impact, which we did not do in 2012 and which we will also report on here. In addition, we reviewed many similar empathy building efforts such as Participatory Design, Design thinking workshop, and Usability research design, where user feedback is integrated into the design process. We took those inspirations and simplified it into a lightweight and streamlined process with minimum time commitment, so that anyone without prior UX knowledge could participate.

The Pokerface Program

Pokerface requires about a three-hour commitment from each Googler participant. This includes participating in a seminar on the basics of research and moderating a user session, the study session itself, and the debrief. In addition to this time commitment, we (the researchers) met with team leads in order to understand their questions and prepare a study script that team members would follow during their sessions, as described below.

Preparing for Pokerface

The first step was for individual teams to meet with a researcher to discuss the teams' research questions. Each researcher helps a team form and refine their questions, and creates a draft user study script to address the ques-

tions. This script is then refined in collaboration with the team and the researcher.

Pokerface: The day of

After scripts have been prepared, participants take part in a one-hour seminar with a researcher where they learn the basic dos and don'ts of running a usability study and have a lesson in research ethics. Next, teams break off into usability sessions, where we partner two participants with one user; one participant moderates the session, and the other takes notes. Finally, the third hour is spent with participants regrouping to debrief and compare findings among the users [2].

Pokerface logistics

In order to scale the program, we created an internal website, sign-up system, and promotional videos involving all of the executives who endorsed the program. These materials aided in communicating the program and its benefits to Googlers, and in helping us convince them to participate.

To manage execution of Pokerface at a large scale, we templated the entire process, from participant recruitment to training material, study scripts, debriefing techniques, and the final report. Hence, the researchers facilitating each session could adopt the program with less of a time commitment than would otherwise be required.

Challenges for Usability at Scale

After scaling up the original program, we were able to send more than 1500 employees at ten locations worldwide through Pokerface within eight months. We experienced many of the same challenges as the original program but we also found a new set of challenges related to the increased scale:

Skepticism from and relevance of user studies for backend engineering teams:

Since we were targeting the entire organization, this meant that some participants were working on back-end (e.g., database management) or infrastructure (e.g., engineering on data server or system coordination). These teams sometimes did not consider user research to be as meaningful for them as did front-end engineers. Some stated that their users are other engineers who collaborate with them and that they could not see how their work could impact end users without needing to do additional research. To overcome this challenge, we adopted the following strategies:

1. We provided customized study scripts to address data-related questions in addition to usability questions.
2. We tried to pair the front- and back-end teams together to unify the discussion and help show back-end teams how their work influenced the user experience.
3. We asked backend engineers to do video confessions about their learnings after the study, and showcased that in emails and videos to share the benefits of the study with teams that had not yet participated.
4. We had researchers explain usability methods and goals when we designed a study script with the team, so participants are more involved in the study design process and understand how it applies to their team.
5. Our endorsement from the executives and videos showcasing executives' participation boosted confidence in participating in this program even for back-end teams.

Over-generalization of the findings based on anecdotes and unsure about usability study:

Also as a result of our targeting Google organization, and not just those who may have shown a proclivity toward talking to users, most colleague participants were new to the field of user research, and many were unaware of standard procedures. For example why they should talk to a small sample of participants for a usability study or how one can translate data to something meaningful and actionable. Hence, we ensured that the planning meeting and training covered some basic information about user research methods and how one should quantify one's hypotheses or triangulate one's qualitative findings with quantitative results. We also emphasized these points during the debriefing sessions, making sure to note where findings had been supported by prior research or appeared to be outliers.

Time commitment in order to scale worldwide:

This project was no small feat, and it involved 26 of our researchers who had to keep their research programs running while also helping with Pokerface. In an attempt to keep time commitment from individual researchers to a minimum, we created a list of criteria and materials that local researchers could adopt and tailor to their partner teams' product areas and research questions. However, even with these procedures this project ended up taking up much more time than initially budgeted. In response to these time pressures, we centralized script writing (which took up much of the preparation time) and instead clearly defined the roles for individual researchers to include handling day of logistics such as picking up participants and facilitating debriefs. By centralizing script writing, we were also able to draw more on prior scripts when building scripts for future teams. While each team had something specific they wanted to learn, we were able to build up a question library that could help to craft each individual team's script.

In addition, we set up a rolling recruitment system with more general participant recruiting criteria in order to streamline the process. While this had the downfall of not allowing participant teams to specify in detail the type of person they wanted to speak to, it allowed us to continuously recruit for future sessions using the same materials that we used for prior sessions (and taking another time consuming task away from the researchers running sessions).

Outcomes and Impact

Another difference in this round of Pokerface from the 2012–2013 version [2] was our increased focus on evaluating the program's impact. In order to measure impact, we ran 2 sets of surveys: one pre-survey sent to participants to complete upon sign-up for the program (but before starting the program), and one post-survey sent to participants to complete shortly after completing the program. The survey is intended to measure participants' attitudes toward Google's user centered design principles. The Pokerface program is intended to build user empathy among Googlers in hopes that this empathy will translate to more awareness of an adherence to Google's design principles.

The pre-survey was sent to all participants and asked them to rate their agreement (on a 5-point bipolar, ordinal Agreement scale) with a set of questions/constructs.

1. My team's product demonstrates clear user benefits.
2. My team's product is easy to use.
3. My team's product is beautiful and robust.
4. My product team is inspired to create excellent products.
5. My product team captures feedback from users.

6. My product team acts on feedback from users.

The Agreement scale used was: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree. This same set of questions was also asked in the post-survey in order to evaluate whether completing the Pokerface program allows participants build user empathy and be more aware when evaluating their team's work and product in terms of user needs. Due to the fact that this program has not been run since 2013 and that surveys were not previously used to evaluate this type of program, we made sure that the survey questions were tested and piloted extensively through internal survey programs in order to know that they would be sensitive enough to detect changes.

We ran the Pokerface program with more than 1500 Googler participants. In order to avoid bias in our survey responses, we made sure that every participant was aware that the surveys were anonymous and were not part of any team or personal evaluation. We also did not reveal the overall survey results until the teams had completed the program to ensure that the results did not influence individual responses.

Of the >1500 participants, >900 participants completed both pre- and post-surveys. In order to test whether there was a statistically significant change in participants' responses, we ran a paired sample t-test on participants' pre-survey and post-survey responses using data from the participants who responded to both surveys. In order to run this test, we converted the bipolar scale to an ordered, interval scale (1-5). Based on prior testing, we can reasonably assume that the distances between each of the response options in the Agreement scale are equidistant. We report the results of this test in Table 1.

We saw significant (at the 95% level) attitude changes in five out of six survey questions. Pokerface participants reported lower levels of agreement with the following three survey questions: (1) "My team's product demonstrates clear user benefits," (2) "My team's product is easy to use," and (3) "My team's product is beautiful and robust." This is what we anticipated—we wanted participants to have more user empathy and also be aware of user issues firsthand in their own product, inspiring them to be more proactive in fixing these usability issues. Participants also reported higher levels of agreement with the following two survey questions: (5) "My product team captures feedback from users," and (6) "My product team acts on feedback from users." This is what we anticipated as well—Pokerface is about giving non-researchers the opportunity to interact with a user face-to-face and allow them to collect data from these users during the user study, then act on the feedback or plan how to act on the feedback at the debrief session. One survey question showed no change in agreement pre- and post-Pokerface: (4) "My product team is inspired to create excellent products." This was as expected as participants are already doing this.

Although the survey did not explicitly include an "empathy" question, based on the survey results, we saw that by talking to users firsthand, our teams understood that we needed to have better user-centered design like having clearer user benefits and easier-to-use products in order to improve our users' experiences. The resulting impact was a cultural change of realizing how much more user research we could integrate into our development process to truly address users' problems and concerns. We captured that in our qualitative feedback from our post-survey. When we asked participants what they learned from this experience, many participants felt the program was eye-opening and they expressed the value of hearing user stories firsthand:

	<i>My team's product demonstrates clear user benefits.</i>	<i>My team's product is easy to use.</i>	<i>My team's product is beautiful and robust.</i>	<i>My product team is inspired to create excellent products.</i>	<i>My product team captures feedback from users.</i>	<i>My product team acts on feedback from users.</i>
Avg diff in pre- and post-survey	-0.37	-0.39	-0.09	-0.03	0.29	0.12
Std Dev	0.89	0.97	0.86	0.79	1.04	0.98
Std Error	0.03	0.03	0.03	0.03	0.04	0.03
p-value	0.00***	0.00***	0.00***	0.32	0.00***	0.00***

* $p < .05$. ** $p < .01$. *** $p < .001$

Table 1: Results of paired sample t-test.

"Observing and talking to the study participants was eye-opening (and humbling)—we came away with very important findings that will influence our product plans in a big way."

"Cognitive load on the user is a precious resource. I don't think we realized how important it is before watching users in person."

"It's quite valuable to talk to real users and watch them struggle. It puts a human side onto the engineering work that we do, which is sometimes easy to forget about."

"It was a very tangible way to remember that we need to put users first. Awesome!"

Most importantly, we have seen many team members take action based on these initial findings. For instance, one team member stated, *"Seeing someone fail right in front of you really motivates you want to fix that problem immediately."*

Conclusion

Project Pokerface was a Google internal user empathy campaign that encouraged user-involvement as part of the product development cycle. We saw that scaling such a program across a large organization of over 1500 employees across 15 sites brought about challenges, however our evaluation showed success, as the program helped our teams worldwide break preconceived perceptions, develop new perspectives, and build more empathy for users and confidence in user-centered research.

As a team, our next steps are to follow-up on this program's initial successes. After exposing team members more broadly to research, we have been getting many requests to continue the program. This is a logistical challenge due to reasons stated above, so we are working to further refine and streamline the process so we can continue to offer the program with minimal disruption to our other research activities.

Acknowledgements

We'd like to thank Jens Riegelsberger, Elizabeth Reid, Elizabeth Churchill, David Huffaker, and Ben Davison for their careful review and comments. We thank our colleagues for their time and immense contributions to this program: Aaron Cheang, Aidan Bryant, Alessandra Millar, Anastassia Tsivy, Ben Dury, Cara de Ferranti, Charles Lai, Constanze Ehrlich, Emily Chee, Hendrik Mueller, Kimberly Ha, Lamar Alayoubi, Martin Henschke, Miguel Malheiros, Muzayun Mukhtar, Nicholas Araujo, Nick Donovan, Nicole Fernandez, Patti Bao, Rich Truncellito, Robert Beaton, Sophie Barness, Susan Szymiczek, and Valerie Lugo-Roussel.

REFERENCES

1. 2017. Design Sprint Kit. <https://designsprintkit.withgoogle.com/case-studies/the-worldwide-tribe-explores-virtual-empathy/>. (2017).
2. Asif Baki, Patrick Bowen, Brianna Brekke, Ferrall-Nunge, Gueorgi Kossinets, Jens Riegelsberger, Nina Weber, and Marissa Mayer. 2013. Project Pokerface: Building a User-Centered Culture at Scale. In *Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA.
3. P. Bosch-Sijtsema and J. Bosch. 2015. User Involvement throughout the Innovation Process in High-Tech Industries. *Journal of Product Innovation Management* 32, 5 (2015), 793–807.
4. Design Council. 2006. Double diamond design process. <https://www.designcouncil.org.uk/news-opinion/design-process-what-double-diamond>. (2006).
5. W.J. Orlikowski. 2000. Using Technology and Constituting Structures : A Practice Lens for Studying Technology in Organizations. *Organization Science* 11, 4 (2000), 404–428.