Measuring Response Rate and Increasing Satisfaction in Innovative Environments: The Impact of Feedback

Lanen J. Vaughn
Michael J. Bortnick
John Carey
Vince R. Orgovan
Joseph Munko
Microsoft Corporation
Redmond, WA 98052, USA
mbortnic@microsoft.com

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

ACM 978-1-4503-5621-3/18/04…$15.00
https://doi.org/10.1145/3170427.3174356

Abstract
Engineering productivity has been a popular topic of discussion from both an employee and employer perspective. Through rigorous experimentation and multi-methods approach, alongside subjective responses, we saw significant improvements in our engineering system. This paper reflects over a year of learnings on methods to increase response rates and designing surveys for monitoring and impacting both efficiency and satisfaction for internal developers. Measuring these concepts became important as we were making significant changes to our engineering system. These methods can be applied in other contexts to increase response rates and monitor satisfaction overtime.

Author Keywords
Response rates; satisfaction; productivity; organizational psychology; participant gratuity; longitudinal experimentation; feedback;

ACM Classification Keywords
H.5.2. Information interfaces and presentation (e.g., HCI): Measurement; Theory; Experimentation

Introduction
This paper is the results of multiscale efforts across Engineering, User Research, Research, and Data Science to improve the Engineering System. This work was informed by behavioral, attitudinal, observational,

and social systems research. We used these methods to get a holistic view evaluating the productivity [2, 8], efficiency, and satisfaction of the engineering system. The results from this work fall more broadly under productivity in innovative environments and could apply to other similar groups.

We've developed a process for understanding, measuring, and monitoring engineering health. We started this work with one primary motivation: **to improve productivity while maintaining satisfaction among engineers**. We consider productivity not as only an individual metric, rather a collective metric of impact on customer experience; however, an individual's perception of their productivity is important to consider in relation to satisfaction. We hypothesize the relationship between satisfaction and productivity at work to have a sweet spot for productivity and satisfaction. We were aware that we needed to be mindful that we could have high productivity in context of high satisfaction or low satisfaction. This project covered these two possibly conflicting motivations, increasing productivity and satisfaction, and we sought to be sensitive to over indexing on productivity and not making our developers feel like they are on a death march [5]. We hypothesized that there is a given set of activities that achieve the high-satisfaction, high-productivity equilibrium. To optimize these activities, we needed to have a listening system in place to make sure that we are not damaging satisfaction while making a difference in productivity of the system. This paper will cover two aspects of the project: the methods we used to monitor satisfaction and the outcomes resulting from these decisions.

Survey Design: Making the potential responses actionable and grounded

While honing our methods to increase response rates, we were simultaneously working to ensure that the data we did collect would be actionable. This parallel work started with a set of exploratory contextual inquires of 30 developers to which we applied grounded theory [3] to better understand the developer experience. From the contextual inquires we realized there was quite a bit of negative sentiment amongst developers in relation to hardware, culture, user experience of tools, and performance of the services and tools they used. Additionally, by shadowing the developers we determined there are a set of activities that developers engage in and the proportion of time spent in these activities varied by level and position. We specifically noted the kinds of activities the developers engaged in, the amount of time spent in each activity, and their subjective experience of that activity. The activity research was also evaluated with a more specific productivity focus by an adjacent research team [2]. This work informed the categories and questions we asked in our first of many quarterly surveys and our concurrent experimentation. As our contextual inquires informed the survey areas, we could link our survey results to actual activities and sentiment we had observed.

We built our survey questions around the primary question "How satisfied are you with the Engineering System?" and specifically selected sub-areas that we felt contributed to satisfaction with the Engineering System overall based on the contextual inquires.

Survey Response Rate: Getting users to respond

We knew our respondents are busy individuals and that would impact response rates. For any of the work we were doing, we knew that high response rates would be important for reducing our overall margin of error. We hypothesized four reasons our participants may not respond: poor incentives, survey fatigue, bystander effect, and procrastination or forgetfulness. We elected to target these four areas as while we measured satisfaction alongside productivity.

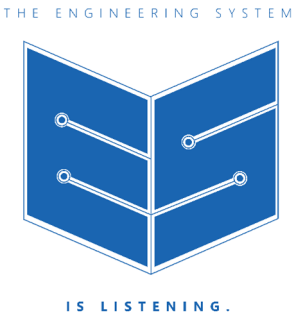

Figure 1: The Engineering System is Listening branding.

We concentrated most of our efforts on ensuring we had the proper incentives to encourage responses from our sample population. Survey requests often don't give potential respondents sufficient *motivation* to respond. Finding the right monetary incentive, or non-monetary incentive, can vary based on the population. Previous research with mail surveys indicated that *pre-payment* had a larger effect on response rates in comparison to other kinds of payment such as raffle or post-hoc payment [1]. Gratuity is only one of several ways to increase incentive, and it is often the least cost-effective way. Monetary incentives can be useful, but not necessarily the best or most effective ones. In fact, in pro-social situations, non-monetary incentives may have better outcomes for participation [4]. We focused heavily on non-monetary incentives to get our developers to respond to our satisfaction survey.

Knowing that we would do longitudinal evaluations, we also targeted reducing survey fatigue. We had a lot of questions we wanted to ask to make sure we were covering all the aspects of satisfaction with the engineering system. Respondents can be tired of taking surveys or get tired of responding during the survey [7]. Survey fatigue can lead to potential respondents not attempting the survey or aborting the survey at partial completion. For some analyses, full competition of a survey is necessary to have the best predictive power. Reducing the fatigue could have an impact on response rate and we employed methods to reduce this concern.

Potential respondents often also believe that others will respond to the survey and they don't need to respond. This belief is called the bystander effect: the belief that someone else will act and therefore the potential respondent is not motivated to act [6]. The bystander effect is compounded by the number of potential bystanders meaning the more people our potential respondents think has been solicited to respond to a survey the less likely they are to respond. Given our survey was for the engineering system there was potential for respondents to believe someone else may respond for them.

Moreover, potential respondents can intend to complete the survey at some later time, but then fail to remember to do it at a more convenient time. Providing reminders to participants to respond, especially when they are likely to have a moment would likely increase response rate.

These areas of attention: respondent motivation, reducing fatigue, individual accountability, and fostering commitment to the survey were crucial to our long-term success in building a feedback system that would foster gains in engineering satisfaction. We used branding techniques, targeted areas of potential non-response bias, model-based survey design, and experimental action across this work.

Methods

In this section we'll cover the efforts we made in branding, invitation and expectations, monetary incentives, and the follow-up to maximize survey responses.

Branding: Prepping our users for our listening system

Before embarking on our complex research plan, we began an internal marketing campaign to prime our developers for the efforts we'd be making. This campaign centered around "the Engineering System is listening" helped to brand the work and later drive response rates for the various experiments. These graphics were used in posters and stickers, and in email communications and on results documentation.

Survey invitations and expectations
In designing our recruitment materials, we targeted the motivation of our potential respondents such as lack of incentive, survey fatigue, the bystander effect, and procrastination. We had 4 incentive-based strategies, 3 which were not monetary: make the survey feel important, make each individual respondent feel important, emphasize the utility of the data, and gratuity.

We strove to make the survey feel important by *conveyed to developers that they would be helping themselves and the greater organization by completing the survey*. The feedback would help the Engineering Systems team determine how to improve the Individual Contributor developer experience. *We also informed each invitee that their manager would be told they were invited*. This made the survey seem more like an official feedback channel and make sure the managers understood we had asked for this time. Moreover, *the invitation to the survey was sent by the Corporate Vice President (CVP) of the Engineering System*. Few developers and their managers ever receive individual emails from a CVP, so it made the invitation stand out, and highlighted the importance of the survey to the CVP.

We strove to make each invitee feel individually important by doing *individual instead of a mass email*. We sent individual email invitations to invitees, instead of a single 'announcement' email with all participants cc'ed/bcc'ed. Furthermore, we *informed that all responses would be read*. Every single response would be read by members of the research team and leadership. *We also informed our potential respondents that very few were invited to the survey.* In our invites we indicated that they were one of a very small set of people invited and this invite-few-strategy only worked if almost everyone responded. *Questions sent to the survey team were responded to within one business hour.* We emphasize the utility of their data and *provided access to the results*. The invitees were informed in the invite that they would get a response from the Engineering System Corporate Vice President about what we learned, and what actions we were taking as a response to the survey. The individual invitations, response, explanation of their importance, and promptly responding to questions, offered evidence to our potential respondents that we recognized their importance.

To decrease the bystander effect, we sought to make the respondent feel that their unique perspective is important to capture, that no one else can or will speak for them. This statement was made more believable as we explained very few people were invited to respond. We affirmed, in the invitation, that we wanted their individual perspective and feedback. We told them that they were one of a very small set of people invited, and that this 'invite few strategy' only worked if almost everyone responds. We *informed invitees that their manager was aware of the invitation.* This made it harder for the invitee to feel as if their silence would not be missed and less able to rationalize letting 'other people' do it. We *requested that invitees who couldn't complete the survey contact us immediately, so we could find a replacement.* Very few invitees asked us to find a replacement, but it created a sense that they were responsible to complete the survey

We had three strategies to mitigate survey fatigue: section gating, random sample without replacement for a year, and setting time expectations. Before each area specific section, such as building code, testing code, etc., of the survey, we asked the respondent a question that determines whether the section applies to them. If not, the section was not presented. Gating was quite important as the survey was quite lengthy (30-35 minutes) and it was often not possible to reduce the survey fatigue very much for everyone. We explained in our invitation that invitees would not be asked to complete another ES Quarterly Feedback survey for at least a year. By randomly sampling a small proportion

of the larger group, we could make this promise. We were also transparent about how long the survey would take. We wanted to avoid drop off in responses where an invitee would start the survey, be surprised by the length, and then quit the survey and not return. Based on the contents of the survey we offered our potential respondents a prediction of the time it would take to complete.

Monetary incentive experimentation
We offered a pre-paid gratuity incentive and experimented with different rewards systems to determine which was the best for response rates. Given that we wanted to reduce the cost of our survey and increase response rates, we experimented with both methods. The two incentive techniques included similar language regarding the importance of their participation and only differed on gratuity.

First, we A/B tested informing invitees of a raffle with $300 (Feb - 16) or $500 (May - 16) for doing the survey. For both raffle surveys, respondents were invited by email. We had nearly the same response rate for both surveys and did not continue the raffle incentive in future surveys. Second, we A/B tested the impact of delivering an individually printed physical 'Thank you' card to invitees, along with a meal card ("we're buying you lunch!") either before or after the survey. We placed the $10 pre-paid gift-cards in a card invitation we left on each developer's desk before they started the work day. 18% more invitees completed the survey when the received the 'Thank you' before the survey instead of receiving afterward.

Following-up with respondents
We sent *reminder emails*, individually, invitees who hadn't completed the survey and warned in our invitation they would receive an email reminding them if they haven't responded. By the Aug-16 survey, we were only sending two reminder emails over the week the survey was open. We also sent the emails first thing in the morning to try to be at the top of the invitee's inbox before they engaged in the day's work. We also had a *slight punishment for procrastination*. The invitees were informed of when they had to complete the survey to avoid the next reminder email. Completing the survey sooner rather than later would reduce the annoyance of reminder emails.

Moreover, we followed up with those respondents having specific concerns. These follow-ups reinforced that the larger team was reading and listening to every response submitted. Every single response was read by our CVP. After analyzing the data, the results were sent to all the engineers by the CVP and he explained what action he would take based on the feedback from the team.

Findings
Survey response rate
The efforts in incentives, reducing fatigue, explaining the importance of their feedback, and the follow-up did not go unnoticed by our respondents. We frequently received feedback regarding the survey, branding, experiments, and feedback we provided about Engineering Systems work. For example, in the last survey one respondent mentioned "It's great to see ES actively seeking feedback The meal card guilt-trip trick worked. :-)" and another echoed our branding "Thank you for listening!" The increase to 100% response rate is attributed to the consistent messaging, the gratuity method, and the repeated feedback from our team to the developers we were serving.

Month-Year	Response rate	Incentive
Nov-15	23%	No incentives
Feb-16	75%	Lottery $300 gift-card
May-16	72%	Lottery $500 gift-card
Aug-16	82%	A/B $10 gift-card pre/post survey
Nov-16	61%	Pre-paid $10 gift-card
Jan-17	90%	Pre-paid $10 gift-card
Apr-17	80%	Pre-paid $10 gift-card
Aug-17	100%	Pre-paid $10 gift-card

Table 1: The relationships between the response rates and incentive methods over time. *Note*: Response rates are typically lower in November due to holidays.

Survey results

Most of these areas of satisfaction with the engineering system that arose from the data aligned with work some teams were already doing. There were a few areas that as they proved to be important to the developers we surveyed, and these areas were given more resources to improve satisfaction. These areas we monitored as we experimented with making changes we hoped would help the productivity in these areas.

For the areas that remained consistent overtime, we monitored the proportions over time and saw while very satisfied remained nearly the same proportion across all surveys, we saw a statistically significant decrease in not satisfied enough ($z=2.83$, $p<0.01$). This decrease in not satisfied enough was echoed in the feedback we would get in the surveys and as we conducted experiments with internal participants. The two major asks that led to an increase in satisfaction and productivity were the distribution of new hardware to those on older machines and the transition of the Engineering System from Source Depot as the code repository to Git. We attribute the decrease in the proportion not satisfied enough to 1) acting on the feedback from our respondents and 2) communicating with our respondents how we acted on all the feedback we received.

Discussion

The results of this case study are valuable in that they pull together some best practices and knowledge from various independent sources. Many research papers on survey methodology typically focus on one method rather than how a collection of methods and practices interact together. We hope that sharing these studies with our broader community that others may find this collection helpful for increasing their response rates and monitoring satisfaction.

What's Next

We had two major challenges with using a survey as our primary listening system to overcome: How do we get people to respond and how do we make their responses actionable? We thought of these two questions as part of a recurring game, where if we could get feedback that was actionable and prove we had acted on their feedback, more developers would be likely to respond in the future. Next, we are planning to repeat these methods across other groups to sharpen the core of the process and test for reliability. We continue to monitor the health of our engineering system and provide feedback to our engineers regarding our actions.

Acknowledgements

This work would not have been possible without the contributions of Umer Farooq, Sam Ladner, Bill Fulton, John Hrvatin, Steven Hicks, and Anthony Kinsey.

References

1. Aljoscha Kaplan, Glenn White. 2002. Incentives in a Business Survey: A Study in Improving Response Rates. *Joint Statistical Meetings- Section on Survey Research Methods:* 1756-1761.

2. André N. Meyer, Thomas Fritz, Gail C. Murphy, and Thomas Zimmermann. 2014. Software developers' perceptions of productivity. In *Proceedings of the 22nd ACM SIGSOFT International Symposium on Foundations of Software Engineering* (FSE 2014). ACM, New York, NY, USA, 19-29. DOI: https://doi.org/10.1145/2635868.2635892
3. Barney G. Glaser, Anselm L. Strauss. 1967. The discovery of grounded theory: strategies for qualitative research. Chicago: Aldine.
4. Dan Ariely, Anat Bracha, Stephan Meier. 2009. Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially. American Economic Review. 99(1): 544-55.
5. Edward Yourdon. 2003. Death March. Yourdon Press.
6. John M. Darley, Bibb Latané. 1968. Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*. 8: 377–383. DOI: https://doi.org/10.1037/h0025589.
7. Paul Lavrakas. 2008. Encyclopedia of survey research methods. London: Sage Publications.
8. Ryan Fuller. 2016. The Paradox of Workplace Productivity. Retrieved April, 28, 2016 from https://hbr.org/2016/04/the-paradox-of-workplace-productivity