From: Andrew Kortina <█████████████████>
To: jeevacation@gmail.com
Subject: Improving Operations Quality through Measurement and Performance Management
Date: Thu, 19 Apr 2018 16:15:52 +0000

Hi ██████,

I'm one of the co-founders of Fin and wanted to let you know about some of the improvements we've made over the past couple months to deliver quality experiences. I also wanted to offer you a $30 credit towards your first requests if you want to try Fin out.

Earlier this year, I wrote about some changes we were making at Fin to level up the quality of our service. This included providing a Dedicated Account Manager for every customer, as a well as a bunch of updates to our internal metrics and processes. In this post, I want to talk a bit more about improvements to service quality that resulted from two of these changes: (i) more comprehensive and accurate measurement of quality signals and (ii) better 'windowing' of metrics to accelerate performance management.

How do you accurately measure service quality?

A few common methods for measuring quality of service for a team include CSAT and NPS. CSAT can be useful for gauging relative quality (either comparing the average performance of different individuals on a team, or comparing your company to another in a similar category). NPS is commonly used to predict customer loyalty and historical trends in the quality of a service or product over time.

Because both of these methods rely on customer surveys and provide only a sampling of customer sentiment and feedback, neither has 100% coverage that alerts you of all poor customer interactions. You must assume that many customers who had a poor experience will not take the time to provide you feedback. Furthermore, sometimes a customer may subjectively label an interaction poor when it's objectively not a failure. (NB: This may be the result of mismatched expectations, which is arguably itself a kind of failure if you subscribe to 'customer is always right' doctrine, but that is outside the scope of this post.)

Completely automated systems (like a sentiment analyzer, eg) may give you broader coverage than user submitted surveys, but also may not be 100% accurate.

Peer review can potentially provide broader sampling than customer surveys, and it may have a more stringent definition of failure that catches some problems that users would miss. But when peers doing the review have team goals that conflict with providing negative feedback to their teammates, you may run into challenges with incentives.

Given that each of these systems for collecting quality signal is imperfect in some way, we have found the best way for scoring quality is to combine and audit all the signals. Here is an overview of how our current quality scoring works:

We want to cast as wide a net as we can and so that we limit our chances of any sort of quality issue escaping opportunities (1) for correction before it negatively impacts a customer and (2) for delivering feedback to the person(s) responsible.

At the same time, since we are managing to quality metrics, we don't want to penalize anyone by decrementing their stats because of a subjective customer response or false alarm of an automated system like our sentiment analysis model. **So, we feed all the various quality signals we collect into a (human) audit process to sift the real issues from the false alarms.** To mitigate challenges of peer review that I mentioned above, we have created a new dedicated Quality Team that works with Account Managers to perform this audit.

Improving Performance Management by Tightening Metric Windows

Once we were confident we were identifying as many quality issues as possible with our new metrics system, we set aggressive goals for our operations team to drive these down to an acceptable rate.

The most impactful change we made to help the team move towards their goals was tightening time windows for key metrics.

As most startups are probably familiar with, it's critical to choose the right time window for each metric: you basically want the smallest time window possible which allows for high enough sample sizes that your metric is relatively stable and free of noise and variance. A tighter time window means you can run experiments to get significant results and feedback faster (and ultimately learn and improve faster). So, daily user counts are better than weekly active counts which are better than monthly active counts, provided your per user activity frequency and total number of active users support the tighter windows.

The same holds for performance metrics for individuals -- tighter windows allow for faster feedback and improvement.

An opportunity we identified early in the quarter was that almost all of our agent performance metrics were pegged to 4 weeks, which meant 4 weeks to identify when someone needs a performance improvement plan, and then 4 weeks to determine the outcome of that plan.

When midway through the quarter we talked about driving results by the end of the quarter, it became clear this feedback cycle was way too long. So for each of our key performance metrics, we asked 'how tight can we make the window for this metric to collect enough data to accurately measure it?' For many important quality metrics, that window was now 2 weeks. That meant for a set of metrics, we could identify the need for a PIP based on only 2 weeks of data, and someone could successfully pass within 2 more weeks. **This doubled the speed of our performance feedback loop, from 8 weeks to 4 weeks.**

It had the additional benefit of empowering each individual on the team to more quickly understand how changes in their workflow, incorporation of coaching, and attention to detail translated into better results, because it is far easier to move a 14 day average metric with a single day of hard work than to move a 28 day average. **Seeing the results of their efforts reflected in key metrics more quickly was a big psychological boost to the team.**

Takeaways

This chart visualizes these effects over the quarter. You can see our Quality Issue Rate spike up early in the quarter, when we launched more comprehensive peer review systems and CSAT email surveys, capturing signal that previously went unmeasured. Then, the Issue Rate steadily marched down as we made improvements to our performance management process:

Given the breadth and complexity of the work Fin does, we've found very few one size fits all answers. This holds both for various methods of measuring quality, as well as for time windows for different key metrics. Ultimately, ensuring the success of our customers hinges upon our ability to measure and drive the quality and efficiency of operations, so we are constantly on the lookout for new, more accurate, and more comprehensive opportunities for measurement.

Thanks,

Andrew Kortina

If you're interested, you can signup using this link to get a $30 credit towards your first requests.

P.S. If this isn't of interest let us know, and we will not reach out in the future with promotional emails like this!