

How DDI's research became faster, cheaper and more reliable



DDI is a global consultancy firm that specializes in assessing and developing frontline, mid-level, and executive-level leaders. They use research to create and validate science-based solutions for every leadership skill.



The task

In 2019, DDI started developing a comprehensive suite of scales to measure leaders' personality, motivations, and experiences.

For example, to create a single scale that measures Ambition, DDI would design about 30 questions and test them on a participant pool of line reports and supervisors in the US and the UK. Any questions that fail to relate to Ambition are removed, while other questions are refined and re-tested until the scale is proven to work well.

DDI researchers also check how each scale responds to reliability analysis, factor analysis, and item response theory analysis, as well as whether each scale correlates in the expected range with already validated scales measuring the same concept.

In the end, including iterations, each scale would require around 500-600 participants.



The challenge

It can be hard to predict whether testing will result in a new reliable scale, but a scale can usually be calibrated with enough iterations. However, fewer iterations mean the research is conducted faster, more efficiently and costs less. Too many iterations means it becomes too expensive and might lead to contradictory findings from the different data collections.

For more than 10 years, the DDI research team had been using Amazon Mechanical Turk (MTurk), a crowdsourcing marketplace for outsourcing tasks to online "workers". Whilst it was a cost-effective solution in the beginning, around 2019 they'd noticed that data quality and speed had started to deteriorate.

To find the right participants in the US, you could pay for the filters in Mturk, but sometimes it would take days to fill the study and responses were clearly less attentive and reliable.

Projects were lasting longer than expected, and DDI increasingly had to discard unreliable responses. This resulted in more required iterations which in turn increased costs.



The solution

DDI wanted a more trustworthy and reliable online platform that could provide hundreds of US line reports and supervisors on demand.

After hearing recommendations about Prolific's high quality participants from PhD students and published research articles, DDI decided to trial Prolific.

"We figured Prolific's people might be a little bit more conscientious and we needed that."

One of the first things that stood out were Prolific's free audience filters. Rather than having to pay extra for each audience selection, DDI could select people who live and work in the US, then filter to line reports or supervisors.

"Prolific has amazing filters. These filters in mTurk are expensive. On Prolific, it's free. And with each filter I apply it tells me how many active people Prolific has, so I know if there are enough and can predict how long the study might take."

Speed of results was also another major benefit, meaning deadlines were more reliably met.

"Also, do you know that Prolific is so fast? I don't know how it does it but sometimes I have my data in two hours! Even with some filters which have lower participant numbers, I know I'll get my sample within a day. With MTurk, it sometimes takes three days to get results."

The biggest difference was felt in the data quality.

"Prolific's workers are more responsive and engaged. The platform makes approving their responses and communicating with them easy too. Now we've completely transferred to Prolific."



The results

15% improvement in data quality

On MTurk, about 15% and sometimes up to 20% of responses had to be discarded. On Prolific, this reduced to just under 5%.

Data gathered in a third of the time

Data collection ranges from 2 hours to 12 hours and rarely beyond, depending on how large the required sample size is and what time of the day the study is published.

10% cheaper

Fewer iterations due to better data quality and free audience filters meant money was saved, despite Prolific requiring participants to be paid more.

"Participants deserve to be paid fairly and we appreciate that. We see it in the quality of their responses. Because MTurk charges for each additional filter added and Prolific gives you access to all filters, we estimate that we save at least 10% every time we use Prolific over MTurk." (Georgi Yankov, Ph.D., Senior Research Scientist)

A boost in production performance

Scales are produced faster, and time is money! With MTurk's clunky design and idiosyncratic participant pool, sometimes half a day is spent in project management and that is about the time an experienced psychometrician takes to analyze thoroughly about five scales.

Sign up to research

[Get started](#)