Data Collection Reflections

Group 5: Ahmet Furkan Ün, Enis Mert Kuzu, Taner Giray Sönmez

An MTurk Crisis? Shifts in Data Quality and the Impact on Study Results:

The paper presents evidence of a decrease in data quality on Amazon's Mechanical Turk (MTurk) platform for psychological research, beginning in the summer of 2018. The study uses a four-wave naturalistic experimental design to show increases in participants' failing response validity indicators, decreases in reliability and validity of a widely used personality measure, and failures to replicate well-established findings. However, the authors suggest that these negative effects can be mitigated by using response validity indicators and screening the data.

Researchers conclude by stating that although MTurk has been an important resource for psychological science, there is compelling evidence of a decrease in data quality which can have a substantial negative impact on study results and conclusions. They suggest that researchers should be aware of this trend and take steps to ensure the quality of their data.

The data collection method used in this study has several good data collection practices that were used to ensure the reliability and accuracy of the study results. One key practice was to use an identical survey across all waves (phases) of the study, which helped to ensure that any changes in results were not caused by variations in the questions being asked. By using the same survey in every wave, the researchers were able to directly compare the results and identify any patterns or trends that emerged over time.

Another important good practice was to collect large amounts of high-quality data. This helped to increase the reliability and accuracy of the study results, as a larger sample size and higher quality data can help to reduce the impact of random error and measurement error. The study also used a specific recruitment method, Mechanical Turk (MTurk) which is useful for replication and understanding bias.

To further understand the impact of different data collection methods, the researchers conducted an additional wave (Wave 4a) using different requirements (HIT) than the previous waves. This allowed them to compare the results from the different waves and understand how changes in the requirements (HIT) might have affected the data quality and study results. Overall, these good data collection practices helped to ensure that the study results were reliable, accurate, and robust and that any patterns or trends that emerged over time could be confidently attributed to real differences in the data.

Researchers also use several bad data collection practices that could have negatively impacted the reliability and validity of the study results. One such practice was using Mechanical Turk (MTurk) to find participants. While this can be an efficient way to collect data, it can also limit the scope or generalizability of the findings, as the sample may not be representative of the general population.

Another bad practice was setting high standards for participants, such as requiring a high approval rate. This can lead to a less diverse and less representative sample, as certain groups of people may be more likely to meet these standards. Additionally, it's not reported whether the researchers used any measures to control for potential sources of bias in the sample, such as

demographic characteristics. This lack of control may result in a sample that is not representative of the population.

In conclusion, using Mechanical Turk (MTurk) alone to find participants and setting high standards for participants can limit the scope(generalizability) of the findings and make the sample less diverse and less representative of the general population. Additionally, not controlling for potential sources of bias can result in a sample that is not representative of the population, which can lead to inaccurate conclusions and unreliable results.

Coloring in the Links: Capturing Social Ties as They are Perceived

The paper explores the limitations of current methods for modeling relationships in network science. The authors note that relationships are typically represented simply as links, perhaps with weights, and that this lack of finer granularity is due in part to the fact that, aside from linkage and strength, no fundamental or immediately obvious dimensions exist along which to categorize relationships.

To address this issue, the authors propose a new set of dimensions that capture major components of many relationships. These dimensions are derived from relevant academic literature and people's everyday descriptions of their relationships. The authors first review prominent findings in sociology and social psychology, highlighting dimensions that have been widely used to categorize social relationships. They then examine the validity of these dimensions empirically in two crowd-sourced experiments. Ultimately, they arrive at a set of ten major dimensions that can be used to categorize relationships: similarity, trust, romance, social support, identity, respect, knowledge exchange, power, fun, and conflict.

The authors argue that these ten dimensions offer higher resolution than existing models, and that one can more accurately predict missing links in a social graph by using these dimensions than by using a state-of-the-art link embeddedness method. They also introduce tinghy.org, an online platform that collects data about how social media users perceive their online relationships. This platform allows them to examine these dimensions at scale. Overall, the authors' goal is to propose a new way of modeling social graphs, which will contribute to theory in network science and practice in designing social-networking applications.

The data collection method used in this research has several strengths that can enhance the generalizability and validity of the findings. Firstly, the researchers used a web-based survey to collect data from participants, which allows for many participants to be recruited and data to be collected from them in a relatively short amount of time. This method can provide researchers with a cost-effective and efficient way of collecting data from a large sample of participants.

Additionally, the researchers used open-ended questions in their survey, which allows participants to describe their relationships in their own words. This approach can help to capture the complexity and diversity of people's perceptions and descriptions of their relationships. By allowing participants to express themselves freely, open-ended questions can provide more in-depth and nuanced information about the topic being studied.Furthermore, the researchers used Mechanical Turk (MTurk) to recruit a large sample of participants, which can help to increase the generalizability of the findings. MTurk is an online platform that allows researchers to access a large and diverse pool of participants quickly and efficiently.

Lastly, the researchers eliminated redundant words in the data collection, which can help to increase the relevance and coherence of the concepts and clusters that emerged from the data.

This step can help the researcher to focus on the most important and meaningful information in the data, and thus to draw more valid and reliable conclusions.

In conclusion, the data collection method used in this research has several strengths that can enhance the generalizability and validity of the findings. The use of web-based survey, open-ended questions, Mechanical Turk, and elimination of the redundant words in the data collection process can provide a more comprehensive and accurate understanding of the phenomenon being studied.

The data collection method used in this research has several limitations that can affect the generalizability and validity of the findings. Firstly, the participants were only from Canada, United Kingdom, and United States, which limits the ability to apply the findings to other cultures and regions. It is important to consider that cultural and regional differences may affect the results, and therefore, the generalizability of the findings is limited.

Additionally, the researchers used a snowball sampling method where they only included papers that were referenced in and referencing the seed papers. This approach may introduce a bias as papers that are not referenced by these papers may have been omitted from the review. This bias is called "publication bias" and it occurs when studies with positive or statistically significant results are more likely to be published than studies with negative or insignificant results. This can lead to an overestimation of the true effect size of the phenomenon being studied.

Moreover, the researchers did not report any measures they took to ensure the quality of the articles included in the review. This can limit the validity of the findings and make it difficult to trust the conclusions drawn from the data. Ensuring the quality of the articles included in a systematic review is a crucial step in the process, as it helps to minimize the risk of including biased or low-quality studies. This can be done by assessing the study's methodology, sample size, and data analysis plan.

As a result, the research's data collection method has several limitations that can affect the generalizability and validity of the findings. It is important for the researchers to be transparent about these limitations and consider their impact on the conclusions drawn from the study.