Can we trust students to collect data ready to use in scientific research?

December 29 2023

Army mentors students and educators. Credit: U.S. Army Research Laboratory

Recent efforts to improve on the openness and transparency in science have already begun paying off towards greater integrity in the way researchers do and report science. It is now common practice for

scientists to pre-register their studies and to openly share their materials and data so that their research is easily available to scientific scrutiny and collaborations.

However, behaviors during [data](#) [collection](#) are still somewhat of a "black box," especially when done by students. In fact, there are plenty of questionable and even fraudulent behaviors, such as telling participants the specific hypotheses of interest before starting the study or even instructing them to answer in a certain way, which are almost impossible to detect later.

The major problem is that current practices and regulations are mostly ineffective in preventing or checking for problematic behaviors in the data collection process. Furthermore, detailed knowledge about the prevalence of such behaviors is relatively scarce. Past research has predominantly focused on questionable practices and misconduct in other stages of the research process like data analysis and reporting.

Importantly, questionable or even fraudulent behaviors might not only be problematic among researchers but also highly relevant in student projects. If public, the data collected from students may be reused by other students, supervisors and other researchers as part of their own work, including [research articles](#) published in journals. Yet, there is no way for those reusers to fully be aware of what has been going during the data collection.

This is how a joint team of psychology students and researchers from LMU Munich decided to investigate students' questionable practices and research misconduct during data collection.

"We wondered: Can we trust student data?" says Dr. Marlene Altenmüller, corresponding author of the article.

"We wanted to know whether and how students actually engage in questionable and even fraudulent practices when collecting data for their projects. And, we were interested in situational factors potentially amplifying or alleviating students' engagement in such behaviors," she further explains.

The research team surveyed 473 psychology students and 199 supervisors at German-speaking universities. They asked them about 17 behaviors, ranging from questionable to fraudulent, to determine whether and which of those the students had engaged in in previous projects. Examples include knowingly letting participants take part in the study while being aware they know the hypotheses; taking part in one's own survey; and deleting or creating data from scratch.

The researchers also sought to assess the students' experiences during their projects. For example, they inquired what kind of expectations and future data use their supervisor had communicated to them. The research team then also asked supervisors about their perceptions of students' data collection behaviors and what they thought about how their students experienced their projects and their supervision.

The survey results reveal some reassuring, as well as some troubling insights into "the black box" of student data collection. While 64% of students did not report any problematic data collection practices, some behaviors were not uncommon: 4% admitted to having deleted data; 8% had participated in their own study; and 26% had let participants take part despite them knowing the hypothesis.

The work is published in the journal *Social Psychological Bulletin*.

On average, supervisors had similar impressions of students' questionable and fraudulent behaviors. Among the notable differences were that supervisors assumed a much lower prevalence for students to

have taken part in their own survey and a lower prevalence for them to have deleted data.

Thus, the research team concluded, supervisors might be underestimating some highly problematic behaviors among students.

To reduce the prevalence of problematic data collection behaviors among students, and therefore prompt better quality of data, the researchers recommend addressing students' perceptions of pressures, opportunities and rationalizations for engaging in these behaviors. Additionally, it would be helpful to make Open Science a central element of teaching.

Particularly, transparent and clear communication between students and supervisors might be one of the most important keys to high-quality, research-ready student data. Those students who knew their data would be used by others, also reported lower prevalence rates of problematic behaviors.

"Supervisors should maybe consider how empirical student projects may not only be an opportunity for teaching, but also for research," conclude the authors.

Provided by The Polish Association of Social Psychology