Expression of Concern: The Decoy Effect as a Nudge: Boosting Hand Hygiene With a Worse Option

The first two authors of the study by Li, Sun, and Chen (2018; Meng Li and Yan Sun) requested that an Expression of Concern be published regarding their article because there is some uncertainty regarding the data underlying their reports. As Editor in Chief of Psychological Science, I have investigated this matter, including consultations with a Psychological Science Statistical Advisor, and have decided to issue an Expression of Concern.

One ground for concern is that the researchers themselves were not directly involved in obtaining or overseeing the measures, which were collected in food-processing factories. According to the authors, a factory manager created spreadsheets with the data collected by quality control workers and sent the spreadsheets to Chen via a mobile message app called “Wechat.”

Unfortunately, the original Wechat messages from the manager were reportedly lost, and efforts to recover them were unsuccessful. Chen shared Excel files that she reported exactly copy what she received from the managers and sent to the other two authors during the time the studies were conducted (2014 for Experiments 1 and 2, 2017 for Experiment 3), but the chain of custody of the data was far from ideal.

The uncertainty regarding the provenance of the data is heightened by the presence of some peculiarities in the data. These anomalies were brought to the authors’ and the editor’s attention by Leif Nelson, Frank Yu, Uri Simonson, and two anonymous researchers. I thank these individuals for bringing their concerns to my attention and for giving me time to investigate the matter.

The first concern is that Li et al. reported that the scales used in Experiment 1 were accurate to 5 grams. Consistent with that, all of the sanitizer-usage data from the intervention period of Experiment 1 contained last digits of 0 or 5. But during the baseline period, 17% of the last digits were not divisible by 5, and 14.9% were 3s and 7s. Li and Sun speculated that different quality control workers used scales of different accuracy levels to record the end weight at the end of a day and the beginning weight the next morning. They reported that on the 2018 online marketplace in China, the brand of scale used in Experiment 1 has adjustable accuracy levels, including 1 gram, 2 grams, and 5 grams, but they could not ascertain the accuracy levels available for the specific scales used in Experiment 1 in 2014. The authors acknowledged that the oddity in last-digit distribution could also be due to other factors, such as human error.

Another peculiarity in the data is that in Experiment 3, the distribution of last digits in the sanitizer-usage data was not uniform. For example, the digits 0 to 4 each occurred at frequencies ranging from 4.1% to 7.1%, while the digits 5 to 9 occurred at frequencies ranging from 11.8% to 15.2%. The authors acknowledged that they do not have an explanation for this nonuniformity.

A final concern is that in Experiment 2, the daily mean sanitizer use in the control group and experimental group were too similar during the baseline period (see Days 1–20 in Fig. 3b of the original article), given the modest size of the groups (20 in each) and given the considerable variability from person to person and from day to day within groups (see Fig. 1 here, which was derived from an analysis done by the first author, Li). The first two authors reported that they were unable to ascertain the source of this unlikely similarity in total sanitizer usage between the two rooms during the baseline period. Day-to-day fluctuations in sanitizer usage were likely partly due to variations in production quantity, which would have had parallel effects on both rooms, but still it is odd that the means for the two rooms were so very similar during the baseline period.

These considerations undermine confidence in these data. But, in my opinion, they do not constitute clear evidence of fraud. I also note that Li and Sun cooperated very helpfully in the investigation of this case.

—D. Stephen Lindsay

Reference
Fig. 1. Each participant’s sanitizer usage on each of the 40 days of Experiment 2, separated by group. Each color represents a different worker. The intervention started on Day 21.