

Supplemental Materials

Analyses conducted by Wansink et al. (2005) not pre-registered by our team

Like Wansink et al. (2005), we found no significant differences in responses to the following statements: “I carefully monitored how much soup I ate,” “I always try to clean my plate (or bowl) at home,” “If other people keep eating, I am more likely to also,” “Eating with other people distracted me from how much I was eating,” and to the following questions: “How full are you right now?,” and “How much food do you think you could eat right now?”

Correlations

Mirroring Wansink and colleagues’ results, we found no significant correlation between responses to “I carefully paid attention to how much I ate” and actual consumption volume. However, unlike Wansink et al. (2005), we found a significant positive correlation between responses to “I carefully monitored how much soup I ate” and actual consumption volume in the experimental condition, $r(216) = .14, p = .034$. Unlike Wansink et al. (2005), we also found a significant and positive association between responses to “If other people keep eating, I am more likely to also” and actual consumption volume in the control condition only, $r(244) = .16, p = .011$. Lastly, unlike Wansink et al. (2005), we found a significant and positive association between responses to “Eating with other people distracted me from how much I was eating” and actual consumption volume in the control condition only, $r(240) = .16, p = .015$. Although these correlations are interesting, these analyses were not pre-registered, and given potential alpha accumulation, we caution against undue interpretation.

Distractor Items

The following questions were included to maintain the cover story and therefore were not included in the analyses. “What is your favorite color?” “Please describe the types of bowls you

normally use at home (color, size, etc.),” “Do you think the color of food indicates how it will taste?” “What is your favorite food, and what color(s) is it?” and “What is a food you don’t like, and what color(s) is it?” all open-ended, and “How tasty is the soup?” and “How much did you enjoy eating the soup?” (both on a 1-5 Likert scale).

Exclusion Criteria

Participants were excluded from the analyses for one or more of the following reasons: If participants saw the refilling tube underneath the table, if participants moved the bowls and saw the holes in the table, if participants were familiar with the original Wansink et al. study, or if participants noticed their soup was refilling in their bowls. If any of the participants in a group found out the purpose of the study through the aforementioned reason(s), we also excluded others in the same group unless the participant explicitly stated that they did not tell the other participants about the purpose of the study.

We found that the results held when including the 168 participants in analyses. That is, participants in the experimental group consumed significantly more soup than did those in the control group, $F(1, 616) = 32.71, p < .001, d = 0.47$. Moreover, we found the same results when running the analyses in only the excluded 168 participants. Participants in the experimental group consumed significantly more soup than did those in the control group, $F(1, 152) = 9.53, p = .002, d = 0.54$.

Study Setup

We have included all 74 photos we have in a folder accessible at this link:
<https://drive.google.com/drive/folders/1TB5eslHb8p-HT0iAbTi6s6PMsTFjH3wP>

Kevin Nguyen, director of the Psychology Department's shop unit, created schematics so that other researchers can engage in their own replications. The picture of the schematic is also available at the link above.

Cost Breakdown

- \$257 table
- \$298 tech services labor + materials so far (tubing, bolts, pots, bowls)
- \$174.86 tech services build curtain around table
- \$240.35 + \$289.95 plastic thing
- \$120.64 additional tubing
- \$487.20 soup for 100 participants (\$300 for soup + shipping)
- \$506 supplies from Amazon (can opener, aprons, 1 induction plate, plastic containers, spoons, dish racks, food scale, dish drying mat, 4 chairs, pyrex cups, 4 pots, ladles, table cloth)
- **Decorations**
 - \$52 chalkboard stickers
 - \$25 fake plants & map
 - \$100
- **Cash participants**
 - \$500
- **Additional soup**
 - \$753.52
 - \$252.78
 - \$616.78

- \$616.78
- \$29.12: spatulas, drying mat, zip ties
- \$14 thermometer
- \$14 dish towels

Total: \$5,347.98

More photographs of the apparatus can be found at <https://osf.io/8q647/>

Differences in analysis procedures between the initial submission and the revision

We present the following description of differences in analyses between the initial submission and the revision in response to a reviewer's suggestion that we do so: "That is, I would like to ask the authors to consider producing a further supplement in which they describe to readers the difference in their analysis procedures between the initial submission and the revision... I think that such a description of the journey that the authors have been on to make their analyses more robust ... could be a useful reference for teachers of methods."

Due to errors in the initial dataset used to run our analyses, we implemented several measures of re-entering data in our revision to ensure improved rigor in our analysis procedures and increased transparency.

Water intake is a variable that we measured but did not previously report on in the initial submission. While Wansink et al. did not mention water intake in their manuscript, through piloting several participants requested water and we realized it would be necessary to provide it. Our protocol stated to only provide water if participants asked and since water intake was not standardized across participants, we thought it prudent to measure water intake in ounces. In the revision, we ran the main analyses controlling for water intake and found that the results remained the same, $F(1, 461) = 24.15, p < .001$.

Mirroring Wansink and colleagues' study, we included several distractor items to bolster the cover story. While we did not mention these items in the initial submission or include them in the analyses, we now include them in supplemental materials. One example item is "Do you think the color of food indicates how it will taste?"

The following errors arose due to suboptimal data entry procedures. Data entry from paper is rife with opportunity for mistakes. In retrospect, we should have engaged in double data entry from the beginning, rather than the entry-and-audit system we put into place where one research assistant checked another research assistant's work. Moreover, research assistants were responsible for hand-calculating variables that could and should have been calculated via programming in the statistical software. This was the first study in a while in our lab that used paper forms, which we opted for because of the large quantities of hot soup being handled and we worried about computers/iPads getting dirty or spilled on. Therefore, there was a loss of institutional knowledge in the lab about the most rigorous way to engage in paper data entry. The senior author (Tomiyama) should have had closer eyes on the data entry manual and procedures, and she takes responsibility and apologizes for these errors.

Two variables corresponded to the condition that participants were assigned to in our original dataset. "Condition#" was the number of the bowl within the 4 bowls in each table. If "Condition#" was 1 or 3, the participant was in the control condition. If "Condition#" was 2 or 4, the participant was in the experimental condition. The "Condition" variable was simply coded as either 0 or 1, with 0 indicating that the participant was in the control condition and 1 indicating that the participant was in the experimental condition. Due to inconsistencies in conditions between the two variables, we have undergone a careful review of the dataset and used

programming to convert “Condition#” to “Condition.” We have also manually confirmed the accuracy of conditions in the dataset.

We additionally engaged in data quality control processes by randomly selecting 25% of the original dataset and having 2 research assistants re-enter that paper data *de novo*. We set an *a priori* threshold of complete data re-entry if there was more than a 5% discrepancy between the newly entered data and the dataset used in our initial analyses. We did not reach this threshold, but we did reconcile discrepancies between the two newly re-entered datasets (1.1%).

In the dataset used in our initial analyses, we mistakenly had duplicate variables that were supposed to contain the same data. For example, “ExpSoupAte” and “E_Postsoup” were both labeled with “How much soup the participant in the experimental condition ate.” However, the data in these variables differed for some participants and were not fully correlated, as they should have been. To address these discrepancies, we started from the beginning and used programming to calculate the amount of soup eaten, while deleting any excess variables that were not used in the analyses to avoid confusion. We have also thoroughly checked the dataset to fix any participant-level errors.

We have now created a [data dictionary](#) with descriptions of each variable in our dataset. We note that we have deleted excess variables in the posted dataset that were not originally used in the analyses or reported in the initial manuscript. For instance, we previously had a variable named “RA_Audit” in our dataset but have now deleted it. Other variables that we have deleted from the revised dataset include “Shouldbeentered,” “FinalParticipants,” “RA_Assess,” “Condition#,” “RA_Enter,” “E_postsoup,” “C_postsoup,” and “F_waterintake.”