Replication of

Thought for Food: Imagined Consumption Reduces Actual Consumption

by Morewedge, C. K. / Huh, Y. E. / Vosgerau, J. (2010)
in: Science, 330, pp. 1530–1533

Replication Authors:
Felix Holzmeister, Jürgen Huber, Michael Kirchler, and Julia Rose

---

The findings of the study by Morewedge et al. (2010) suggest that habituation to a food item can occur even when its consumption is merely imagined, implying a decrease in one's responsiveness to the food and the motivation to obtain it. Five experiments show that people who repeatedly imagine eating a food many times subsequently consume less of the imagined food than people who repeatedly imagine eating that food fewer times, imagine eating a different food, or do not imagine eating a food. These results suggest that mental representation alone can engender habituation to a stimulus.

Hypothesis to replicate and bet on:

Repeatedly imagining eating a food subsequently reduces the actual consumption of that food (a comparison of the 30-repetition treatment and the control treatment in experiment 1; independent samples t-test, $t(30) = 2.78$, $p = 0.0092$, provided by the original authors. The analysis in the original study pools the variance across the 30-repetition, the 3-repetition, and the control condition and reports an ANOVA result of $F(1, 46) = 4.50$, $p = 0.0393$, p. 1531.)

(This hypothesis was picked by lottery instead of comparing the mean consumption of M&M's between the 30-repetition treatment and the 3-repetition treatment; $F(1, 46) = 5.81$, $p < 0.05$, p. 1531).

Power Analysis and Criteria for Replication: First Data Collection

The original sample size is 32 participants and the standardized effect size measured as the correlation coefficient (r) is 0.453. To have 90% power to detect 75% of the original effect size, a sample size of 89 is required. The criteria for replication are an effect in the same direction as the original study and a p-value < 0.05 (in a two-sided test).

Power Analysis and Criteria for Replication: Second Data Collection

If the original result is not replicated in the first data collection, a second data collection is carried out. To have 90% power to detect 50% of the original effect size in the pooled sample (first and second data collection), a sample size of 211 is required, i.e., a sample size of 122 in the second data collection is required. The criteria for replication are an effect in the same direction as in the original study and

a p-value < 0.05 (in a two-sided test) in the pooled data.

Sample

The sample in the first data collection consists of 89 students from the University of Innsbruck. If the original result is not replicated in the first data collection (two-sided p-value < 0.05 in the same direction as the original study), a second data collection consisting of 122 additional students from the University of Innsbruck will be carried out such that the pooled sample size is 211. There are no exclusion criteria.

Materials

We use the original instructions, which have been made available by the authors, and pictures resembling the ones used in the original study together with M&M's as the food to be (imagined to be) consumed. While the original study has been conducted using Adobe Authorware which is no longer supported, the replication experiments will be programmed in z-Tree (Fischbacher, 2007). As the replication study is conducted in German, all materials from the original study are translated from English to German.

Procedure

We follow the procedure of the original study, with only slight but unavoidable deviations as outlined below. The following summary of the experimental procedure is therefore based on the section "Method" (pp. 2–3) of the Supporting Online Material.

Similar to the original study, participants in the replication experiment are run in groups ranging in size from 3 to 9 participants. Each participant completes the experiment in a private cubicle with at least one empty cubicle separating participants in order to ensure that participants cannot see or hear other participants eating during the experimental session. At the beginning of the experiment, all participants indicate how much they like M&M's on a seven-point scale (Dislike Extremely (1) and Like Extremely (7)) and how often they use coin-operated parkometers on a five-point scale (Never (1) and Weekly (5)).

Next, all participants imagine performing 33 repetitive actions, which serve as the critical manipulation. The actions participants imagine are manipulated in a between-subjects design. Participants in the control condition are shown a picture of a small white bowl filled with 33 20-cent coins. They are instructed to imagine inserting those 33 coins into a parkometer one-at-a-time.

Participants in a 30-repetition condition are first shown a picture of a small white bowl filled with three 20-cent coins and imagine inserting those three coins into a parkometer, one-at-a-time. After completing this task, these participants are shown a picture of a small white bowl containing 30 M&M's and are asked to imagine eating those 30 M&M's, one-at-a-time.

To ensure that participants actually engage in the imagination task that they are instructed to perform, the instructions stress that participants should try to imagine the actions exactly as described. After finishing the imagination task, participants estimate the actual size of an M&M and a 20-cent coin by indicating which of five images of an M&M and which of five images of a 20-cent coin best represent their actual size.

Participants are then told that they would perform a taste-test. They are each given a small bowl containing 40g of M&M's, and are encouraged to take their time and eat as many M&M's as they would like to before answering questions about the M&M's. Participants notify the experimenter when they finish eating; she then removes the bowl of M&M's

(which is weighed after the participant left the laboratory). After the tasting stage, participants answer filler questions about M&M's (e.g., "how sweet are they?", "what is your favorite M&M's color?") and report how they felt right then on the PANAS, a measure of positive and negative affect. Finally, participants list when and what they last ate before the experiment and are compensated.

Analysis

The analysis will slightly deviate from the analysis reported in the original study. The comparison of consumed M&M's in the 30-repetition and control treatment reported in the original study is based on an ANOVA including a third treatment as well (3-repetition condition), pooling the variance across all three conditions. In the replication study, only the 30-repetition and control condition will be replicated. Thus, rather than replicating the ANOVA result reported in the original study, the two-sample comparison (independent samples t-test; not reported in the original study) between the consumed amount of M&M's in the 30-repetition and control condition will be replicated after prior agreement with the original authors.

In the original study, participants, on average, consumed 4.08 grams ($SD = 0.33$) of M&M's in the control condition compared to 2.21 grams ($SD = 0.48$) in the 30-repetition condition. Based on the result of an independent samples t-test provided by the original authors, the difference in the consumed amount of M&M's between the two conditions is statistically significant with a p-value of 0.0092 ($t(30) = 2.78$). The same test will be used in the replication study. As in the original study, observations more extreme than 2.5 standard deviations from the overall mean will be dropped from the analysis.

The results will first be estimated based on the first data collection. If the original result is replicated in the first data collection (a two-sided p-value < 0.05 in the same direction as in the original study), the second data collection will not be carried out. If the original result is not replicated in the first data collection a second data collection will be carried out. The above statistical test will then be estimated for the pooled sample of the first and second data collection to test if the original result replicated (a two-sided p-value < 0.05 in the same direction as in the original study).

Differences from Original Study

The replication procedure is identical to that of the original study, with some unavoidable deviations. The replication will be performed at the University of Innsbruck between September 2016 and September 2017, while the original data was gathered at Carnegie Mellon University in 2008. The experiment will be conducted in German rather than in English (as the original study).

Participants in the original study received course credits to participate in the experiment. In contrast, a monetary show-up fee for participating in the experiment will be used in the replication. At the Innsbruck EconLab it is common practice to compensate participants with about €15.00 per hour (on average) for participating in experiments, with a minimum payment of €5.00 for participating in an experiment. Participation in the original study, on average, took about 15 minutes. In order to comply with the terms of use of the Innsbruck EconLab, subjects will be paid a show-up fee of €5.00 for participating in the replication experiment.

The original study contains five experiments testing whether repeated mental simulation of experiencing a stimulus can engender habituation: for the replication, the focus is only on a comparison of the actual food consumption in the 30-repetition treatment and the control treatment in experiment 1.

While the result in the original study refers to an ANOVA pooling the variance across three treatments, only the two conditions of interest and the two-sample comparison (independent samples t-test) will be replicated after prior agreement with the original authors.

In the original study, participants are instructed to imagine inserting the coins into a coin-operated laundry machine. As coin-operated laundry machines – most likely – are not as common in Austria as in the United States, the term "laundry machine" is substituted by "parkometer" in the replication experiments, according to prior agreement with the original authors. While sessions in the original experiment were run in groups ranging in size from 1 to 4 participants, sessions in the replication experiments will be run in groups of 3 to 9. Though, as in the original experiment, at least one empty cubicle will separate participants to ensure that participants cannot see or hear other participants eating during the session.

Replication Results for the First Data Collection (90% power to detect 75% of the original effect size)

The main result for the replication experiment is calculated as in the original study. In the original article, participants, on average, consumed 4.08 grams ($SD = 0.33$) of M&M's in the control condition compared to 2.21 grams ($SD = 0.48$) in the 30-repetition condition ($t(30) = 2.78$; $p = 0.0092$).

For the first stage data collection, a sample of 96 participants has been collected. As in the original study, observations more extreme than 2.5 standard deviations from the overall mean were dropped from the analysis. The overall mean of consumed M&M's was $M = 10.29$ grams ($SD = 8.99$). 6 participants consumed more than $M + 2.5SD = 32.76$ grams of M&M's and one participant did not consume any M&M's at all such that 7 observations were excluded from the analysis. Accordingly, the sample for the replication result is $N = 89$ as planned.

In the replication experiment, participants in the control condition ($N = 44$), on average, consumed 10.30 grams ($SD = 5.28$) of M&M's; participants in the 30-repetition condition ($N = 45$), on average consumed 6.69 grams ($SD = 4.30$). Based on an independent samples t-test, the difference in the amount of consumed M&M's between the two conditions is statistically significant with a p-value of 0.0006 ($t(87) = 3.5384$). The original result is thus replicated.

For completeness, we also report the result without excluding the observation with a consumption of 0 grams of M&Ms as this additional exclusion criteria has not been addressed in the pre-replication version of the replication report. Participants in the control condition on average consumed 10.07 grams ($SD = 5.44$) while participants, on average, consumed 6.69 grams ($SD = 4.30$) in the 30-repetition condition ($t(88) = 3.2686$; $p = 0.0015$). As above, the difference between the two conditions is statistically significant and in the same direction as in the original study.

Replication Results for the First and Second Data Collection Pooled (90% power to detect 50% of the original effect size)

Since the original result has been replicated in the first data collection with 90% power to detect 75% of the original effect size with a statistically significant effect ($p < 0.05$) in the same direction as in the original study, no second data collection has been conducted.

Unplanned Protocol Deviations

The replication experiments were conducted exactly in the way as described above,

except for the additional exclusion criteria regarding participants refusing to consume any M&Ms at all. Since no participant refused to consume any M&Ms in the original study, the original authors did not have to consider this exclusion. However, this exclusion criteria was decided on together with the original authors before starting the data collection.

Discussion

Given the criteria and procedure outlined above, the hypothesis of interest has been replicated at a significance level of $\alpha < 5\%$ ($p = 0.0006$) with a power $\pi = 90\%$ to detect 75% of the original effect size.

References

Fischbacher, U. (2007): *z-Tree: Zurich Toolbox for Ready-Made Economic Experiments*, Experimental Economics, 10(2), pp. 171–178.

Morewedge, C. K. / Huh, Y. E. / Vosgerau, J. (2010): *Thought for Food: Imagined Consumption Reduces Actual Consumption*, Science, 330, pp. 1530–1533.