AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

2025, Vol. 154, No. 11, 2935–2943
https://doi.org/10.1037/xge0000964

Arbitrary Stimuli Are Not Devalued by Stopping Action: A Registered Replication of Wessel et al. (2014)

Patrick G. Bissett[1], Kriti G. Achyutuni[1], Jamie K. Li[1], Henry M. Jones[2], Sunjae Shim[3], Jaime Ali H. Rios[4], Logan J. Bennett[1], Margot Mitchell[1], Jules Jackson[1], Maya Bar Or[5, 6], Dana Roll[5], Naama Meidan[5, 6], Anastasia Saliy-Grigoryan[5], Tehya M. LePage Drummond[7], Shriya S. Pokala[7], Jessica R. Cohen[7], Tom Schonberg[5, 6], and Russell A. Poldrack[1]

[1] Department of Psychology, Stanford University
[2] Department of Psychology, The University of Chicago
[3] Department of Neuroscience, University of California, Berkeley
[4] Department of Psychology and Neuroscience, Duke University
[5] School of Neurobiology, Biochemistry and Biophysics, Faculty of Life Sciences, Tel Aviv University
[6] Sagol School of Neuroscience, Tel Aviv University
[7] Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill

There is a sizable literature on inhibition-induced devaluation (IID), in which withholding a behavioral response in the context of a specific stimulus reduces the value of that stimulus and reduces the likelihood of choosing or using that stimulus in the future. IID holds great promise for behavior change and improving public health, because a variety of public health problems, including unhealthy eating and substance abuse, involve the inability to control the use of something of inherent, perceived, or associated value. Wessel et al. (2014) showed that when arbitrary stimuli were associated with response inhibition, subjects valued those stimuli less, suggesting that IID may be possible across a variety of stimuli. In three preregistered direct replications across three universities, including a U.S. private university, a U.S. public university, and a public university in Israel, we show no evidence for IID. This brings into question the robustness of IID as implemented by Wessel et al. using arbitrary stimuli and the stop-signal paradigm. This work encourages future efforts to directly replicate previous findings of IID in related designs, including those using stimuli with inherent value like food.

Public Significance Statement

A primary determinant of public health is individual decisions like whether to eat unhealthy foods or whether to use and abuse drugs and alcohol. Given this, there is a pressing need for robust interventions that can influence or change decisions and actions. One such promising intervention, inhibition-induced devaluation (IID), involves a computer-based procedure in which subjects repeatedly make fast keyboard responses to certain visual stimuli (e.g., healthy foods) but inhibit their responses to others (e.g., unhealthy foods). Previous research has shown that after undergoing this intervention, subjects chose and valued the inhibited stimuli less. In three samples across three universities and two countries, we found no evidence for IID, bringing into question the robustness and efficacy of this intervention. Our work attempted to replicate one specific procedure for IID (Wessel et al., 2014), so we encourage future replication efforts that evaluate the efficacy of other IID procedures.

Keywords: response inhibition, replication, value, behavior change, stop-signal paradigm

Supplemental materials: https://doi.org/10.1037/xge0000964.supp

This article was published Online First August 28, 2025.

Sarah Brown-Schmidt served as action editor.

Patrick G. Bissett https://orcid.org/0000-0003-0854-9404

Code and data to replicate these results are available at https://doi.org/10.5281/zenodo.15548949. Hypotheses and analyses were preregistered at https://osf.io/bt5f3/files/osfstorage. There has been no prior dissemination of the ideas or data that appear in this article other than the following Open Science Framework repositories at https://osf.io/x38aj/ and https://osf.io/kq5xd.

This article was funded by National Institute of Drug Abuse, National Institutes of Health (Grant F32DA041773) awarded to Patrick G. Bissett. The authors thank Jan Wessel for his generous assistance in the preparation of this article, including providing code, sharing data, and extensive guidance on implementing the instructions and procedure for Wessel et al.'s (2014) Experiment 2. The authors also thank Audrey Yin, Jaemon Jumpawong, Emily Kessler, Elizabeth Kudriatsev, Julia Shaw, and Trey Thurman for their assistance running participants.

Large language models, including ChatGPT and Google Gemini, were used to assist with code writing, but all code was evaluated and tested by one or more of the coauthors. Large language models were not used for writing any of the text of this article.

Patrick G. Bissett played a lead role in conceptualization, data curation, funding acquisition, methodology, project administration, supervision, writing–original

continued

Unhealthy behaviors, such as smoking, drug and alcohol abuse, overeating, and a sedentary lifestyle, account for a substantial proportion of the risk associated with preventable premature deaths in the United States (Nielsen et al., 2018). This has led to a pressing need for interventions that result in healthy behavior changes. One promising intervention is inhibition-induced devaluation (IID), in which withholding a behavioral response in the context of a specific stimulus reduces the value of that stimulus and even results in an individual being less willing to choose or use that stimulus in the future. IID holds great promise for improving public health, because a variety of public health problems can be distilled to the inability to control the use of something of inherent or perceived value. The promise of IID is that in devaluing a stimulus (e.g., a palatable but unhealthy food) by associating it with inhibition, an individual will be able to more effectively control their urge to use the valuable but harmful substance, facilitating better choices.

Inhibition is a fundamental mechanism of control that facilitates goal-directed behavior (Logan, 1985). In order to navigate a dynamic world, actions and goals must have the flexibility to stop and change. In such instances, response inhibition is engaged (Bissett & Logan, 2011, 2013; Camalier et al., 2007; Logan & Cowan, 1984). Two main paradigms used to capture response inhibition are the go/no-go task (Donders, 1868/1969) and the stop-signal task (Logan & Cowan, 1984). Both tasks involve the presentation of a series of stimuli, with one or more stimuli indicating the need to respond and a different set of stimuli indicating the need to withhold a response. The go/no-go task involves presenting only one stimulus on any given trial, and subjects make a response (i.e., go) if that stimulus has a certain identity but do not make a response (i.e., no-go) if that stimulus has a different identity. In contrast, the stop-signal task involves making a choice response to a “go” stimulus that is presented on every trial, except for the subset of trials in which a second, subsequent “stop signal” is presented.

There may be fundamental differences between the processes involved in the stop-signal task and the go/no-go task. Computational modeling work has shown that go/no-go data can be well described in terms of a two-choice diffusion process without a separate inhibition component (Gomez et al., 2007). Convergent empirical work has questioned whether go/no-go tasks require response inhibition, particularly when go and stop stimuli have similar trial probabilities and when trials have longer durations (Wessel, 2018). Other work has argued that both tasks involve inhibition but different kinds: cancellation in the stop-signal task and restraint in the go/no-go task (Schachar et al., 2007). Recent work leveraging encephalography and electromyography (Raud et al., 2020) argues that go/no-go and stop-signal tasks recruit different neural mechanisms, with go/no-go mechanisms more akin to response selection than response inhibition, in line with Gomez et al. (2007). Therefore, existing work suggests that the go/no-go task may not involve inhibition or may involve a different kind of inhibition.

Both the stop-signal and the go/no-go task have been used to implement IID. In the canonical implementation, certain stimuli are often or always presented in the context of withholding a motor response, and other stimuli are never presented in the context of withholding a motor response. In a subsequent phase of the study, subjects are either asked to do some form of value judgment on each class of stimuli, given the opportunity to consume the two classes of stimuli (e.g., if the classes are food), or both.

The existing IID literature suggests that, in general, stimuli that were associated with withholding a motor response are valued less and chosen less than stimuli not associated with withholding a motor response. Associating specific food stimuli with no-going (Veling et al., 2013) or stopping (Lawrence et al., 2015) reduces subsequent consumption of the trained food stimuli. Similarly, associating specific alcohol stimuli with no-going (Houben et al., 2011) or stopping (Jones & Field, 2013) reduces subsequent consumption of the trained alcohol stimuli. Meta-analyses of food and alcohol stimuli (Jones et al., 2016) and food stimuli alone (Yang et al., 2019) provide evidence for IID. A multiexperiment preregistered study (Chen et al., 2016) suggests that IID is more effective when inhibition trials are rare. More generally, preregistered and well-powered work suggests that IID using stimuli with inherent value is robust (Chen et al., 2016, 2018; Quandt et al., 2019). However, other work urges some caution on the efficacy of IID for behavior change (Schonberg et al., 2014; Smith et al., 2017).

Additionally, some recent work suggests that devaluation and behavior change may be particularly potent when using a go/no-go instead of a stop-signal task (Allom et al., 2016; Chen et al., 2019; Jones et al., 2016; Yang et al., 2019). This may be because the go/no-go task tends to involve a more consistent mapping between no-go stimuli and withholding a response (i.e., stop success rates tend to be much higher in the canonical implementations of the go/no-go task compared to the stop-signal task), and the effect may be driven by developing a consistent association between specific

draft, and writing–review and editing and a supporting role in formal analysis. Kriti G. Achyutuni played a lead role in formal analysis, software, validation, and visualization and a supporting role in data curation, investigation, writing–original draft, and writing–review and editing. Jamie K. Li played a lead role in investigation and a supporting role in data curation, formal analysis, methodology, software, and writing–review and editing. Henry M. Jones played a supporting role in investigation, software, and writing–review and editing. Sunjae Shim played a supporting role in investigation, software, and writing–review and editing. Jaime Ali H. Rios played a supporting role in investigation, software, and writing–review and editing. Logan J. Bennett played a supporting role in investigation, software, validation, and writing–review and editing. Margot Mitchell played a supporting role in data curation, software, validation, and writing–review and editing. Jules Jackson played a supporting role in investigation and writing–review and editing. Maya Bar Or played a supporting role in investigation and writing–review and editing. Dana Roll played a supporting role in investigation and writing–review and editing. Naama Meidan played a supporting role in investigation and writing–review and editing. Anastasia Saliy-Grigoryan played a supporting role in investigation and writing–review and editing. Tehya M. LePage Drummond played a supporting role in investigation and writing–review and editing. Shriya S. Pokala played a supporting role in investigation and writing–review and editing. Jessica R. Cohen played a supporting role in project administration, resources, supervision, and writing–review and editing. Tom Schonberg played a supporting role in conceptualization, project administration, resources, supervision, and writing–review and editing. Russell A. Poldrack played a lead role in resources and a supporting role in conceptualization, funding acquisition, project administration, supervision, and writing–review and editing.

Correspondence concerning this article should be addressed to Patrick G. Bissett, Department of Psychology, Stanford University, Building 420, 450 Jane Stanford Way, Stanford, CA 94536, United States. Email: pbissett@stanford.edu

stimuli and withholding a response, not between specific stimuli and attempting to withhold a response.

Perhaps most excitingly, work by Wessel et al. (2014) showed that if arbitrary stimuli (colored shapes) are associated with specific monetary values, then a subset of these arbitrary stimuli are associated with stop-signal inhibition, subjects selectively devalue them. This study is exciting because it suggests that IID may not only apply to stimuli with an a priori value (e.g., palatable foods) but may be an effective strategy to experimentally devalue any type of stimuli. When a stimulus is associated with inhibition, it appears to become subjectively less valuable and therefore less likely to be used and (possibly) abused. Additionally, the authors used a stop-signal task and did not have a consistent mapping between stop stimuli and withholding a response, suggesting that IID may be robust to nonconsistent mappings between stimuli and withholding a response. Wessel et al. (2015) replicated this finding in subsequent work and showed that IID was particularly large in individuals who had explicit knowledge of the learned monetary values of the stimuli.

Shortly after this work by Wessel et al. (2014, 2015), our group aimed to replicate and extend this work. In two direct replications of Wessel et al.'s (2014) Experiment 2, we were unable to find evidence of IID. However, these direct replications had relatively small sample sizes and were acquired from only one population of conveniently sampled subjects in the Stanford University area (see Supplement for Direct Replications 1 and 2). We also completed a series of seven conceptual replications that are detailed here (https://osf.io/bt5f3/files/osfstorage) that were unable to find evidence of IID. In order to more robustly evaluate IID, we submitted a registered replication to directly replicate the IID effect with three large, highly powered samples from three different university settings: Stanford University, University of North Carolina at Chapel Hill (UNC), and Tel Aviv University. All samples showed strikingly consistent results: There was no evidence for IID. Equivalence testing suggested that stop and no-stop stimuli had equivalent values, and Bayes factor (BF) analyses showed considerable evidence in favor of the null hypothesis of no IID. We discuss the implications of this null finding for IID of arbitrary stimuli and stimuli with inherent value. All data and code are openly available on the Open Science Framework (https://doi.org/10.5281/zenodo.15548949).

Method

Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulation, and all measures in this article or in the preregistered replication article at https://osf.io/bt5f3/files/osfstorage. All data, analysis code, and research materials are available at https://doi.org/10.5281/zenodo.15548949. Hypotheses and analyses were preregistered at https://osf.io/bt5f3/files/osfstorage, and all procedures followed what was specified in the preregistered replication article unless otherwise specified.

Subjects

Subjects were recruited from three locations: Stanford University, Tel Aviv University, and UNC.

Stanford University Replication (N = 156)

One hundred fifty-six subjects (96 female, 56 male, four declined to state) were recruited from the Stanford Psychology Department Sona Paid subject pool for a single 90-min testing session. One subject was removed due to technical difficulties during the experiment. Subjects were paid \$12 + the reward from five randomly selected trials from Phase 1 (see the Experimental Task section for an explanation of the phases), ranging from \$1 to \$13. In order to facilitate recruitment and to align with more up-to-date standards at Stanford for subject payment, after 89 subjects completed the experiment, the payment range was altered. Subjects after this point were paid a base rate of \$20 and received a reward ranging from \$1 to \$13. The M_{age} was 25.7 (±7.9) years. Ninety-seven of the 156 subjects were Stanford University students. Subjects were excluded if they took part in any of our previous IID studies. This work was approved by the Stanford Institutional Review Board.

Tel Aviv University Replication (N = 151)

One hundred fifty-one subjects (110 female, 40 male, one declined to state) were recruited from the Tel Aviv Sona Paid subject pool, Facebook groups, and ads around campus for a single 90-min testing session. One subject was removed for leaving early due to difficulty concentrating. Payment was translated to new Israeli Shekels (NIS) but aimed to be similar to the U.S. sites and Wessel et al.'s (2014) Experiment 2. Specifically, subjects were paid 40 NIS (approximately \$11) + the reward from five randomly selected trials from Phase 1, ranging from 5 to 41 NIS (approximately between \$1 and \$11). The NIS-to-dollar exchange rate varied over time, with the total compensation being as high as the equivalent of approximately \$25. The M_{age} was 27.7 (±6.3) years. Eighty-eight of the 151 subjects were Tel Aviv University students. This research was approved by the Tel Aviv University ethics committee.

UNC Replication (N = 163)

One hundred sixty-three subjects (113 female, 43 male, six nonbinary, one declined to state) were recruited from Research for Me and flyers placed around campus. Two subjects were removed for failing to complete the study. Subjects were paid a base rate of \$12 + the reward from five randomly selected trials from Phase 1, ranging from \$1.67 to \$16.51. The M_{age} was 25.2 (±8) years. No data were recorded on how many subjects were students at UNC. This work was approved by the UNC Institutional Review Board.

Power Analysis and Exclusion Criteria

In an a priori power analysis, we assumed a small effect size (Cohen's $d = .2$) and power of .95 for the main effect comparing the auction value of stop and no-stop stimuli (see code at https://github.com/bissettp/PowerAnalysisIID), which indicated a minimum of $N = 79$ subjects. This main effect was the primary evidence for IID in Wessel et al. (2014) with a significantly lower auction value for stimuli associated with stopping interpreted as evidence for IID. Thus, we aimed to acquire 79 subjects who satisfied the following exclusion criteria, following Wessel et al., from each of our three independent sites (see https://osf.io/x38aj/ for exclusion scripts): (a) stop-failure reaction times (RTs) ≥ no-stop-signal RT, (b) any reportable explicit knowledge in Phase 2, (c) stop-signal reaction time (SSRT) < 100 ms, (d) the absolute value of the main effect of stopping devaluation >1.5 × the interquartile range (IQR). However, as specified in our preregistration, instead of calculating mean SSRT as in Wessel et al., we calculated integration SSRT with replacement, as suggested by a recent consensus guide for the stop-signal task (Verbruggen et al., 2019). Additionally, when computing SSRT, we used the RTs to stop shapes as the underlying go distribution on stop trials. We found that RTs tended to be slower in response to stop than no-stop shapes, so no-stop-signal RTs on only stop shapes should act as a better, more specific estimate of the underlying go distribution on stop trials.

Experimental Task

All three sites followed the same procedure, except materials were provided in Hebrew at Tel Aviv University where appropriate. The experiment was run with the Psychophysics Toolbox in Matlab. In order to eliminate any coding discrepancies, all three sites used the following code provided by Jan Wessel that was used for Experiment 2 of the original Wessel et al.'s (2014) sample: https://github.com/janwessel/stopdeval. Tones were presented through headphones at Stanford and using computer speakers in a testing room at Tel Aviv and UNC. The experimental task is summarized below, paraphrasing the task description in Wessel et al.

Learning Phase (Phase 1)

In this phase, participants learned to associate eight geometric shapes with four distinct monetary values. The shapes included a square, circle, diamond, triangle, inverted triangle, cross, hexagon, and "I" shape, each colored differently (white, green, blue, yellow, cyan, magenta, orange, or gray) to enhance implicit value associations. Two shapes each were tied to mean values of $0.50, $1.00, $2.00, or $4.00 for U.S. participants and 1.625 NIS, 3.25 NIS, 6.5 NIS, or 13 NIS for the Tel Aviv sample. The distributions had a uniform dispersion of ±25c or 81.25 agorot (100 agorot = 1 NIS) around the means. Shapes were randomly assigned colors and values for each participant.

Each trial began with a fixation cross for 500 ms, followed by one shape appearing in one of four screen quadrants. Participants had 1,000 ms to respond based on the quadrant the stimulus appeared in, or a "too slow" message would display for 1,000 ms. Upon responding, a black square displayed the message "You won $X.XX" for 1,000 ms. Participants were informed that quick and accurate responses would yield higher rewards and that the total payout would come from five randomly selected trials at the end. Trials without responses (misses) or incorrect responses (errors) were repeated until the subject responded correctly. Importantly, the monetary rewards were predetermined and not based on performance, with each shape assigned to one of four fixed payout distributions. Only 80% of trials for each shape had rewards according to these distributions, while the remaining 20% had no reward. Each shape appeared with equal probability in all quadrants, and participants completed 50 trials per shape (400 trials total) across four blocks.

Treatment Phase (Phase 2)

In this phase, some shapes from the learning phase were associated with stopping, and others were not. Participants made quick responses based on shape placement (left or right) and had to stop their response upon hearing a stop signal. Four shapes (one per value-step, i.e., one shape each associated with a mean value of $0.5, $1, $2, or $4 or 1.625 NIS, 3.25 NIS, 6.5 NIS, or 13 NIS) were always paired with go trials (i.e., no-stop signal). The other four were paired with stop signals on 76.4% of trials to avoid making the association between shape and stopping overly predictable and to discourage participants from consciously withholding responses upon seeing shapes associated with stopping.

Each trial displayed the shapes on either side of a vertical line for 500 ms, and participants responded within 1,000 ms using two designated keyboard buttons. Stop signals (200 ms sine-wave tones at 900 Hz) occurred shortly after stimulus onset, prompting participants to cancel their responses. The stop-signal delay was adjusted based on performance to aim for a success rate of 50% (Levitt, 1971). The initial stop-signal delay value was 250 ms and was computed separately for left and right responses depending on ongoing performance (+50 ms following successful stop trials, −50 ms following failed stop trials). Participants completed 34 trials per shape (272 overall) in four blocks and received feedback on their reaction times and error rates during breaks. Additionally, the experimenter would check the blockwise feedback to ensure that RTs were between 400 and 650 ms, the probability of responding given a stop signal was close to .5 (.4–.6), and stop-signal delay was >100 ms. If subjects diverged from these ranges, usually by slowing their responses, they were encouraged at the breaks to adjust their behavior accordingly.

Valuation Phase (Phase 3)

In this phase, participants bid on the shapes seen in the previous phases using an auction procedure to assess their subjective values. Participants chose from six different cent or agorot amounts using specific buttons on the keyboard. They were advised to bid optimally, as both overbidding and underbidding were discouraged.

Each trial began with a fixation cross for 500 ms, followed by the shape being displayed for 1,500 ms. Participants then had 5,000 ms to select a bid amount from six options. Each shape was shown 10 times, resulting in 80 total trials. The bid amounts were drawn from various predetermined sets to encourage diverse bidding behavior. These bid sets were selected to encompass the full range of true values associated with each shape from the learning phase, which included amounts of 40, 80, 160, and 320 cents or 130, 260, 520, and 1,040 agorot, respectively. These values were 80% of the mean values associated with each shape to account for 20% of trials for each shape having reward values of 0. These bids came from five different sets of values, in cents, as specified by Wessel et al. (2014): [34, 68, 102, 136, 170, 204], [39, 78, 117, 156, 195, 234], [44, 88, 132, 176, 220, 264], [49, 98, 147, 196, 245, 294], and [54, 108, 162, 216, 270, 324]. In the Tel Aviv sample, the bids came from these sets of values, in agorot: [135, 270, 405, 540, 675, 810], [154, 308, 462, 616, 770, 924], [116, 232, 348, 464, 580, 696], [173, 346, 519, 692, 865, 1,038], and [192, 384, 576, 768, 960, 1,152]. Each of these sets was displayed twice for each shape, in random order. This approach ensured that participants saw the complete range of values for each shape, presented in random order within four blocks.

Procedure

Participants were instructed briefly on the learning phase and treatment phase and then practiced 10 trials each of both phases. Next, they completed both testing phases. Following this, the experimenter gave participants a sheet with the auction instructions and ensured that participants understood the instructions before participants completed the auction phase. Subjects were then asked six debriefing questions to determine their explicit knowledge in the learning phase and treatment phase. In consultation with Jan Wessel, we deviated from Wessel et al.'s (2014) procedure by using the updated debriefing questions in the README file of the code in Jan Wessel's GitHub specified above (https://github.com/janwessel/stopdeval). The questions were as follows: (a) Did you notice anything in the RT experiment? (b) Did you think you were better for some symbols than for others? (c) Some symbols were worth more than others. Could you try to name any symbols you thought had particularly high or low values? (d) Did you notice anything in the stop-signal experiment? (e) Did you think some symbols were paired with stopping more often than others? (f) Some symbols were paired with stopping more often. Could you try to name them?

Subjects were then compensated for their participation as described in the Subjects section.

Exclusions From Each Site

In Phase 2, subjects were excluded for having their SSRT < 100 ms and/or having their stop-fail RT ≥ no-stop RT. Subjects were also categorized into three categories: full explicit learners, partial explicit learners, and implicit learners. These descriptions are quoted below from Wessel et al. (2014).

1. Full explicit learners: This designation was given if participants spontaneously verbalized one of the regularities in response to the "open" questions of whether they noticed something about the first or second phase of the experiment (Questions 2 and 5) and displayed significant accurate knowledge about the regularity (Questions 3 and 6). They needed to accurately name the relative values of more than half of the stimuli to be labeled as an explicit learner of the learning regularity (Phase 1). They needed to accurately name all four shapes that were either paired with stopping or never paired with stopping to be labeled as an explicit learner of the stopping regularity (Phase 2).

2. Partial explicit learner: This designation was given if a participant reported a "feeling of knowing" of one of the regularities in response to the "open" questions about the learning and stopping phases (Questions 2 and 5, respectively) and could accurately name at least one shape for which the regularity was true (Questions 3 and 6, respectively). For example, in Phase 1, they could state, "The green square was always paired with high reward" or "The diamond was always paired with low rewards." With regard to Phase 2, they could state, for example, "The blue circle was often paired with stopping" or "The triangle was never paired with stopping."

3. Implicit learner: These participants reported not noticing any regularities in response to any of the questions in the debriefing questionnaire (Wessel et al., 2014, p. 8).

Subjects were excluded if they had full explicit knowledge or partial explicit knowledge of the stopping contingency (Phase 2) in order to exclude participants who realized the purpose of the experiment in the treatment phase. Phase 1 explicit knowledge was not an exclusion criterion, but see the Supplemental Material for analyses that isolate Phase 1 explicit learners.

In Phase 3, subjects were excluded for having their IID effect > 1.5 × IQR (i.e., IID effect >75th percentile of IID effects +1.5 × the IQR of the IID effect or IID effect <25% percentile of IID effects −1.5 × the IQR of the IID effect).

Stanford Exclusion

According to the exclusion criteria highlighted above, we excluded a total of 76 subjects from the Stanford sample, leaving a usable sample of 80 subjects. One subject was excluded for failing to complete the experiment due to technical difficulties. Nine subjects failed both behavioral criteria in Phase 2 and Phase 2 explicit knowledge criteria. Fifteen subjects were excluded for failing only behavioral criteria in Phase 2. Forty-four subjects were excluded for having only Phase 2 explicit knowledge. Seven subjects were excluded for having their IID effect > 1.5 × the IQR. Thus, a total of 76 subjects failed exclusion criteria.

Out of the subjects who failed Phase 2 behavioral criteria, 20 subjects were excluded for SSRT < 100 ms, three subjects had stop-fail RT ≥ no-stop RT, and one subject failed both criteria. Out of the subjects who failed explicit knowledge criteria, 44 subjects had partial explicit knowledge, and nine subjects had full explicit knowledge.

Tel Aviv Exclusion

Sixty-seven subjects were excluded from the Tel Aviv sample, leaving a total of 84 usable subjects. Fifteen subjects failed both behavioral criteria in Phase 2 and explicit knowledge criteria in Phase 2. Seven subjects were excluded for only failing behavioral criteria in Phase 2, and 40 subjects were excluded for only having Phase 2 explicit knowledge. Five subjects were excluded for having their IID effect > 1.5 × the IQR, leading to a total of 67 excluded subjects.

Of the subjects excluded for failing behavior, 11 subjects had SSRT < 100 ms, six subjects had stop-fail RT ≥ no-stop RT, and five subjects failed both criteria. Of the subjects excluded for having Phase 2 explicit knowledge, 41 had partial explicit knowledge, and 14 had full explicit knowledge.

UNC Exclusion

Seventy-six subjects were excluded from the UNC sample, leaving a total of 87 usable subjects. Two subjects were removed for failing to complete the study. Nine subjects failed both the behavior and Phase 2 explicit knowledge criteria. Four subjects failed only the behavioral criteria in Phase 2, and 54 subjects had only Phase 2 explicit knowledge. Seven subjects were excluded for having their IID effect > 1.5 × the interquartile range, leading to a total of 76 excluded subjects.

Of the subjects who failed behavioral criteria, 12 subjects had their SSRT < 100 ms, and one subject had SSRT < 100 ms and stop-fail RT ≥ no-stop RT. Of the Phase 2 exclusions, 47 subjects had partial explicit knowledge, and 16 had full explicit knowledge.

Analytical Choices

Equivalence Tests

Equivalence tests were computed using the two one-sided *t*-test procedure (Lakens et al., 2018) in Python. In line with one of the suggested procedures from Lakens et al. (2018), we decided to use the smallest observed effect size that was statistically significant in a previous study as the equivalence margin. The smallest effect size that was statistically significant in Experiment 2 of Wessel et al. (2014) was 0.24. Thus, we used 0.24 as our equivalence margin for evaluating the difference between auction values of stop and no-stop shapes.

BF

The statistical software JASP (JASP Team, 2024) was used to compute BF. A Bayesian paired-samples *t* test was conducted, selecting BF_{01} as the BF type and the alternative hypothesis as no-stop shape auction values ≠ stop shape auction values.

Results

All analyses were specified in a preregistered, registered replication article unless otherwise noted.

Treatment Phase (Phase 2) Results

The stopping (Phase 2) results are detailed and included in Table 1. In all three locations, stop-failure RT was significantly faster than go RT to all shapes, go RT to shapes associated with stopping, and go RT to shapes not associated with stopping (all $ps < .001$). The probability of responding given a signal was .47 for all three locations, and the SSRT was between 191 and 219 ms.

Valuation Phase (Phase 3) Results

Separately for each of the Stanford, Tel Aviv, and UNC samples, we ran a 2 (stopping: stop vs. no-stop shapes) × 4 (value: 0.5 dollars or 1.625 NIS, 1 dollar or 3.25 NIS, 2 dollars or 6.5 NIS, and 4 dollars or 13 NIS) repeated measures analysis of variance that was modeled after the analysis presented in Figure 1. In each sample, there was a significant main effect of value (all $ps < .005$) and no significant interaction (all $ps > .128$). Critically, there was no main effect of stopping (all $ps > .49$; see Table 2 and Figure 1). There were a similar number of subjects who assigned lower or equal values to stop than no-stop stimuli in all three samples. The significant main effect of value but no main effect of stopping was also verified in control analyses in which we (a) included all subjects (i.e., did not exclude for behavior or Phase 2 explicit knowledge) and (b) included only the subset of subjects who had Phase 1 explicit knowledge. Both control analyses were run for each site separately and combined across sites (see the Supplemental Material). Similar results were also found in two smaller direct replications that preceded these three new samples (see the Supplemental Material). BF results showed substantial evidence in favor of the null hypothesis (all $BF_{01} > 6.6$). Additionally, equivalence tests run on the stopping condition showed that the two stopping groups were statistically equivalent (all $ps < .001$).

The results of the repeated measures analysis of variances from Stanford (left panel), Tel Aviv (middle panel), and UNC (right panel) are presented in Figure 1. We summarize the net devaluation results in Figure 2.

Discussion

IID is a promising behavior change technique that hypothesizes that the value of a stimulus is decreased when it is associated with inhibition. We completed five direct replications of Wessel et al.'s (2014) IID experiment, three of which were preregistered and acquired across three different subject populations, and found no evidence for IID. Stimuli associated with inhibition did not have a lower value, and both equivalence testing and Bayesian analyses showed evidence that stimuli associated with inhibition did not differ in their value from those not associated with inhibition. This result was found in every sample we collected (at Stanford, UNC, and Tel Aviv) and in the combined sample across all three sites and was robust to the use of different exclusion criteria. These results provide strong evidence against the presence of an IID effect using Wessel et al.'s procedure.

Why Did We Fail to Replicate Wessel et al.'s (2014) IID Effect?

Wessel et al. (2014) showed an IID effect in two experiments using the stop-signal task and then replicated this devaluation effect in a subsequent sample (Wessel et al., 2015). However, we found strong evidence against the IID effect in five replication attempts across samples at Stanford, UNC, and Tel Aviv. This discrepancy is surprising, and we offer the following considerations for this discrepancy. One possibility that was suggested by a reviewer after Direct Replications 1 and 2 is that the subjects at Stanford University were idiosyncratic, in part because of the high rate of Phase 2 explicit knowledge. In Direct Replications 1 and 2, 38.7% of subjects had some explicit knowledge of the Phase 2 contingency, compared to 22.2% in Wessel et al.'s Experiment 2. However, in our new, three-site replication, we also observed higher rates of Phase 2 explicit knowledge (34.2% at Stanford, 36.6% at Tel Aviv, and

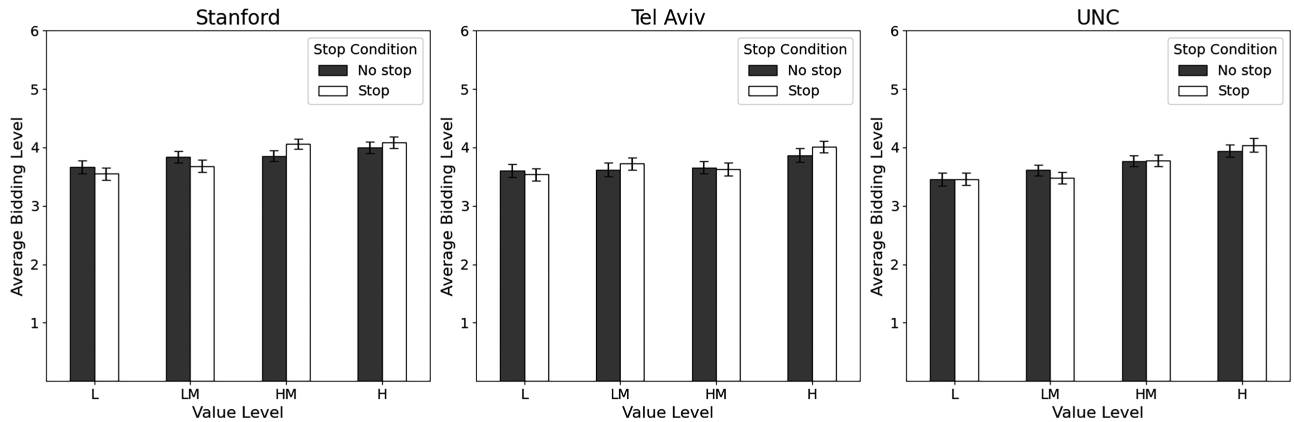
Table 1
Stopping Results From Stanford, Tel Aviv, and UNC

Location	Go RT (ms)	Go RT stop shape (ms)	Go RT nonstop shape (ms)	Stop-failure RT (ms)	p(resp|signal)	SSRT (ms)
Stanford	576 (110)	590 (120)	572 (108)	469 (97)	.47(0.05)	191 (54)
Tel Aviv	619 (123)	632 (134)	616 (123)	506 (107)	.47(0.04)	219 (59)
UNC	583 (85)	600 (98)	579 (85)	478 (83)	.47(0.04)	204 (64)

Note. The mean value is followed by the standard deviation in parentheses. UNC = University of North Carolina at Chapel Hill. RT = reaction time; *p*(resp|signal) = probability of responding given a stop signal; SSRT = stop-signal reaction time.

Figure 1

Auction Results From Stanford (Left), Tel Aviv (Middle), and UNC (Right) Samples



Note. Low (L), low-medium (LM), high-medium (HM), and high (H) signify the four associated value levels from Phase 1 (50 cents or 1.625 new Israeli Shekels [NIS], 1 dollar or 3.25 NIS, 2 dollars or 6.5 NIS, and 4 dollars or 13 NIS, respectively). Mirroring Wessel et al. (2014), error bars indicate the standard error of the mean across subjects. UNC = University of North Carolina at Chapel Hill.

39.1% at UNC), so the preponderance of evidence suggests that this IID procedure results in relatively high rates of Phase 2 explicit knowledge and no IID. Second, though we used the experimental code from Wessel et al. and followed all guidance provided for running the experiment, it is possible that we may have inadvertently diverged from some aspect of the procedure that was key to showing the IID effect. If this is the case, this would suggest that the IID effect observed in Wessel et al. is quite fragile. Third, given the complete lack of an IID effect in this multisite effort that included more subjects across more populations than the original Wessel et al. (2014, 2015) experiments, it is possible that the original Wessel et al. (2014, 2015) IID effect was a false positive that resulted from sampling variability. Relatedly, the original Wessel et al. (2014, 2015) samples included fewer subjects ($N = 27$–65 each) than each of our individual site samples ($N = 151$–163 each), even after applying Wessel et al.'s exclusion criteria ($N = 80$–87 each).

Implication for IID Using Different Designs

Wessel et al.'s (2014, 2015) work consists of two studies among a larger group of IID studies, most of which used stimuli with inherent value, like food (e.g., Veling et al., 2013) or alcohol (Jones et al., 2016), instead of arbitrary stimuli like the colored shapes of Wessel et al. Our work provides strong evidence against the robustness of the IID effect in Wessel et al. and brings into question experiments using stop-signal tasks and arbitrary devaluation stimuli. However, there are multiple large-scale, preregistered studies that support the robustness of IID when using the go/no-go task and stimuli with inherent value. For example, three preregistered samples of 336, 140, and 290 subjects, respectively, observed the IID effect using palatable foods as stimuli and the go/no-go task as the intervention (Adams et al., 2021; Tzavella & Chambers, 2023). Additionally, there are multiple meta-analyses supporting IID using stimuli with inherent value, especially when the intervention is a go/no-go task (Jones et al., 2016; Yang et al., 2019, 2022). Hence, we believe it would be premature to conclude from our work that IID is generally not robust or replicable. However, meta-analyses can be driven by publication bias (Kvarven et al., 2020), and even preregistration can be susceptible to publication bias, as pre-registering a study does not ensure submission or acceptance of the study once it is completed (Ensinck & Lakens, 2025). Only registered replications, like this study, provide an assurance of publication before data collection that is made irrespective of the subsequent results. To our knowledge, this is the first registered replication of the IID effect. To further validate IID using stimuli of inherent value and the go/no-go task as an intervention, we encourage future registered replications of the IID effect, including using the go/no-go task with arbitrary stimuli,

Table 2

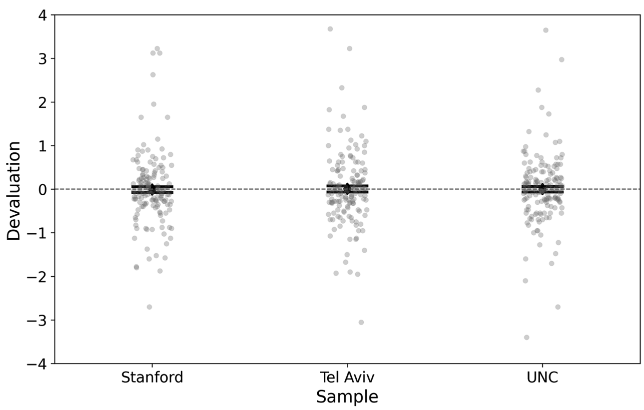
Auction Results for Stanford, Tel Aviv, UNC, and All Sites Combined

Effect	Test	Stanford	Tel Aviv	UNC	Combined
Stopping	Main effect	Stopping $F(1, 79) = 0.01$, $p = .925$	Stopping $F(1, 83) = 0.47$, $p = .495$	Stopping $F(1, 86) = 0.02$, $p = .903$	Stopping $F(1, 250) = 0.21$, $p = .646$
Stopping	BF_{01}	8.08	6.62	8.39	12.74
Stopping	Equivalence test	$t(79) = 4.73$, $p < .001$	$t(83) = 3.60$, $p < .001$	$t(86) = -6.28$, $p < .001$	$t(250) = 8.20$, $p < .001$
Stopping	Count of no stop ≤ stop	46 of 80	46 of 84	44 of 87	136 of 251
Value	Main effect	Value $F(3, 237) = 7.51$, $p < .001$	Value $F(3, 249) = 4.39$, $p = .005$	Value $F(3, 258) = 10.34$, $p < .001$	Value $F(3, 750) = 19.91$, $p < .001$
Interaction	Interaction	Interaction $F(3, 237) = 1.91$, $p = .128$	Interaction $F(3, 249) = 0.63$, $p = .594$	Interaction $F(3, 258) = 0.62$, $p = .602$	Interaction $F(3, 750) = 1.43$, $p = .234$

Note. UNC = University of North Carolina at Chapel Hill; BF = Bayes factor.

Figure 2

Point Plot With 95% Confidence Intervals of the Inhibition-Induced Devaluation Effect for Stanford (Left), Tel Aviv (Middle), and UNC (Right) Subjects



Note. Devaluation is computed by subtracting the mean stop stimulus bidding level from the mean no-stop stimulus bidding level, so values above 0 indicate devaluation. Each dot is an individual subject. UNC = University of North Carolina at Chapel Hill.

the stop-signal task using stimuli with inherent value, and the go/no-go task using stimuli with inherent value.

Why Might Wessel et al.'s (2014) Design Be Less Robust Than Other IID Designs?

A thorough discussion of the theoretical mechanisms underlying IID is beyond the scope of this work, and we point interested readers to the following theoretical work (Saad et al., 2024; Veling et al., 2022), but we briefly discuss two considerations for why the procedure appears less robust than other IID tasks. First, in line with previous work suggesting that the go/no-go task is a better intervention to produce IID than the stop-signal task (Allom et al., 2016; Chen et al., 2019; Jones et al., 2016; Yang et al., 2019), it is possible that the consistent association between stop stimulus and successfully stopping that is a feature of the go/no-go task may be essential to observe IID. Second, Phase 3 of Wessel et al. (2014, 2015) involves an auction to bid money for colored shapes, which subjects may have found more confusing than judging the value of real-world stimuli like food. Our experiments all replicated Wessel et al.'s main effect of value (i.e., subjects offered higher values at the Phase 3 auction for stimuli with higher Phase 1 value), so our procedure was successful in associating stimuli with different values, but the difference in value across stimuli was small (mean auction value across the three replication samples: low = 3.54, *SD* = 0.99; low-medium = 3.66, *SD* = 0.95; high-medium = 3.79, *SD* = 0.91; high = 3.99, *SD* = 0.97; see Figure 2). This small effect may have resulted in part from subjects being uncertain how to assign value to colored shapes.

Constraints on Generality

The present null findings regarding IID are most directly applicable to the specific experimental paradigm of Wessel et al. (2014, Experiment 2), which this multisite registered replication closely followed. While data were collected from three diverse university settings (a U.S. private university, a U.S. public university, and a public university in Israel), broadening the generalizability across academic contexts, the participant samples predominantly consisted of university-affiliated young adults. Consequently, the generalizability of these null findings to populations differing in age or educational background remains to be determined. Additionally, the generalizability of these null findings to tasks using stimuli with inherent value or the go/no-go task as the intervention is uncertain.

Conclusion

In an international three-site preregistered replication, including a U.S. private university, a U.S. public university, and a public university in Israel, we fail to replicate Wessel et al.'s (2014) Experiment 2, bringing into question the robustness of IID effects using arbitrary stimuli like colored shapes and the stop-signal paradigm as the intervention. We encourage similar registered replication efforts to evaluate the robustness of IID effects that implement other designs, including using stimuli with inherent value and the go/no-go task as the intervention. Combined with the results shown here, such future work will help to validate whether the promise of IID for changing behavior and improving public health is supported by a robust effect.

References

Adams, R. C., Button, K. S., Hickey, L., Morrison, S., Smith, A., Bolus, W., Coombs, E., Randolph, S., Hunt, R., Kim, D., Chambers, C. D., & Lawrence, N. S. (2021). Food-related inhibitory control training reduces food liking but not snacking frequency or weight in a large healthy adult sample. *Appetite*, *167*, Article 105601. https://doi.org/10.1016/j.appet.2021.105601

Allom, V., Mullan, B., & Hagger, M. (2016). Does inhibitory control training improve health behaviour? A meta-analysis. *Health Psychology Review*, *10*(2), 168–186. https://doi.org/10.1080/17437199.2015.1051078

Bissett, P. G., & Logan, G. D. (2011). Balancing cognitive demands: Control adjustments in the stop-signal paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(2), 392–404. https://doi.org/10.1037/a0021800

Bissett, P. G., & Logan, G. D. (2013). Stop before you leap: Changing eye and hand movements requires stopping. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(4), 941–946. https://doi.org/10.1037/a0033049

Camalier, C. R., Gotler, A., Murthy, A., Thompson, K. G., Logan, G. D., Palmeri, T. J., & Schall, J. D. (2007). Dynamics of saccade target selection: Race model analysis of double step and search step saccade production in human and macaque. *Vision Research*, *47*(16), 2187–2211. https://doi.org/10.1016/j.visres.2007.04.021

Chen, Z., Holland, R. W., Quandt, J., Dijksterhuis, A., & Veling, H. (2019). When mere action versus inaction leads to robust preference change. *Journal of Personality and Social Psychology*, *117*(4), 721–740. https://doi.org/10.1037/pspa0000158

Chen, Z., Veling, H., Dijksterhuis, A., & Holland, R. W. (2016). How does not responding to appetitive stimuli cause devaluation: Evaluative conditioning or response inhibition? *Journal of Experimental Psychology: General*, *145*(12), 1687–1701. https://doi.org/10.1037/xge0000236

Chen, Z., Veling, H., Dijksterhuis, A., & Holland, R. W. (2018). Do impulsive individuals benefit more from food go/no-go training? Testing the role of inhibition capacity in the no-go devaluation effect. *Appetite*, *124*, 99–110. https://doi.org/10.1016/j.appet.2017.04.024

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica*, *30*, 412–431. https://doi.org/10.1016/0001-6918(69)90065-1 (Originally work published 1868)

Ensinck, E. N. F., & Lakens, D. (2025). An inception-cohort study quantifying how many registered studies are publicly shared. *Advances in Methods and Practices in Psychological Science*, *8*(1), 1–10. https://doi.org/10.1177/25152459241296031

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General*, *136*(3), 389–413. https://doi.org/10.1037/0096-3445.136.3.389

Houben, K., Nederkoorn, C., Wiers, R. W., & Jansen, A. (2011). Resisting temptation: Decreasing alcohol-related affect and drinking behavior by training response inhibition. *Drug and Alcohol Dependence*, *116*(1–3), 132–136. https://doi.org/10.1016/j.drugalcdep.2010.12.011

JASP Team. (2024). *JASP* (Version 0.19.3) [Computer software]. https://jasp-stats.org/

Jones, A., Di Lemma, L. C. G., Robinson, E., Christiansen, P., Nolan, S., Tudur-Smith, C., & Field, M. (2016). Inhibitory control training for appetitive behaviour change: A meta-analytic investigation of mechanisms of action and moderators of effectiveness. *Appetite*, *97*, 16–28. https://doi.org/10.1016/j.appet.2015.11.013

Jones, A., & Field, M. (2013). The effects of cue-specific inhibition training on alcohol consumption in heavy social drinkers. *Experimental and Clinical Psychopharmacology*, *21*(1), 8–16. https://doi.org/10.1037/a0030683

Kvarven, A., Strømland, E., & Johannesson, M. (2020). Comparing meta-analyses and preregistered multiple-laboratory replication projects. *Nature Human Behaviour*, *4*(4), 423–434. https://doi.org/10.1038/s41562-019-0787-z

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269. https://doi.org/10.17605/OSF.IO/QAMC6

Lawrence, N. S., Verbruggen, F., Morrison, S., Adams, R. C., & Chambers, C. D. (2015). Stopping to food can reduce intake: Effects of stimulus-specificity and individual differences in dietary restraint. *Appetite*, *85*, 91–103. https://doi.org/10.1016/j.appet.2014.11.006

Levitt, H. (1971). Transformed up–down methods in psychoacoustics. *Journal of the Acoustical Society of America*, *49*(2), 467–477. https://doi.org/10.1121/1.1912375

Logan, G. D. (1985). Executive control of thought and action. *Acta Psychologica*, *60*(2–3), 193–210. https://doi.org/10.1016/0001-6918(85)90055-1

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, *91*(3), 295–327. https://doi.org/10.1037/0033-295X.91.3.295

Nielsen, L., Riddle, M., King, J. W., Aklin, W. M., Chen, W., Clark, D., Collier, E., Czajkowski, S., Esposito, L., Ferrer, R., Green, P., Hunter, C., Kehl, K., King, R., Onken, L., Simmons, J. M., Stoeckel, L., Stoney, C., Tully, L., … NIH Science of Behavior Change Implementation Team. (2018). The NIH Science of Behavior Change Program: Transforming the science through a focus on mechanisms of change. *Behaviour Research and Therapy*, *101*, 3–11. https://doi.org/10.1016/j.brat.2017.07.002

Quandt, J., Holland, R. W., Chen, Z., & Veling, H. (2019). The role of attention in explaining the no-go devaluation effect: Effects on appetitive food items. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(8), 1119–1133. https://doi.org/10.1037/xhp0000659

Raud, L., Westerhausen, R., Dooley, N., & Huster, R. J. (2020). Differences in unity: The go/no-go and stop signal tasks rely on different mechanisms. *NeuroImage*, *210*, Article 116582. https://doi.org/10.1016/j.neuroimage.2020.116582

Saad, M., Bohon, C., & Weinbach, N. (2024). Mechanisms underlying food devaluation after response inhibition to food. *Appetite*, *199*, 107387. https://doi.org/10.1016/j.appet.2024.107387

Schachar, R., Logan, G. D., Robaey, P., Chen, S., Ickowicz, A., & Barr, C. (2007). Restraint and cancellation: Multiple inhibition deficits in attention deficit hyperactivity disorder. *Journal of Abnormal Child Psychology*, *35*(2), 229–238. https://doi.org/10.1007/s10802-006-9075-2

Schonberg, T., Bakkour, A., Hover, A. M., Mumford, J. A., Nagar, L., Perez, J., & Poldrack, R. A. (2014). Changing value through cued approach: An automatic mechanism of behavior change. *Nature Neuroscience*, *17*(4), 625–630. https://doi.org/10.1038/nn.3673

Smith, J. L., Dash, N. J., Johnstone, S. J., Houben, K., & Field, M. (2017). Current forms of inhibitory training produce no greater reduction in drinking than simple assessment: A preliminary study. *Drug and Alcohol Dependence*, *173*, 47–58. https://doi.org/10.1016/j.drugalcdep.2016.12.018

Tzavella, L., & Chambers, C. D. (2023). Explicit and implicit devaluation effects of food-specific response inhibition training. *Journal of Cognition*, *6*(1), Article 10. https://doi.org/10.5334/joc.256

Veling, H., Aarts, H., & Stroebe, W. (2013). Using stop signals to reduce impulsive choices for palatable unhealthy foods. *British Journal of Health Psychology*, *18*(2), 354–368. https://doi.org/10.1111/j.2044-8287.2012.02092.x

Veling, H., Becker, D., Liu, H., Quandt, J., & Holland, R. W. (2022). How go/no-go training changes behavior: A value-based decision-making perspective. *Current Opinion in Behavioral Sciences*, *47*, 101206. https://doi.org/10.1016/j.cobeha.2022.101206

Verbruggen, F., Aron, A. R., Band, G. P., Beste, C., Bissett, P. G., Brockett, A. T., Brown, J. W., Chamberlain, S. R., Chambers, C. D., Colonius, H., Colzato, L. S., Corneil, B. D., Coxon, J. P., Dupuis, A., Eagle, D. M., Garavan, H., Greenhouse, I., Heathcote, A., Huster, R. J., … Boehler, C. N. (2019). A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. *eLife*, *8*, Article e46323. https://doi.org/10.7554/eLife.46323

Wessel, J. R. (2018). Prepotent motor activity and inhibitory control demands in different variants of the go/no-go paradigm. *Psychophysiology*, *55*(3), Article e12871. https://doi.org/10.1111/psyp.12871

Wessel, J. R., O'Doherty, J. P., Berkebile, M. M., Linderman, D., & Aron, A. R. (2014). Stimulus devaluation induced by stopping action. *Journal of Experimental Psychology: General*, *143*(6), 2316–2329. https://doi.org/10.1037/xge0000022

Wessel, J. R., Tonnesen, A. L., & Aron, A. R. (2015). Stimulus devaluation induced by action stopping is greater for explicit value representations. *Frontiers in Psychology*, *6*, Article 1640. https://doi.org/10.3389/fpsyg.2015.01640

Yang, Y., Qi, L., Morys, F., Wu, Q., & Chen, H. (2022). Food-specific inhibition training for food devaluation: A meta-analysis. *Nutrients*, *14*(7), Article 1363. https://doi.org/10.3390/nu14071363

Yang, Y., Shields, G. S., Wu, Q., Liu, Y., Chen, H., & Guo, C. (2019). Cognitive training on eating behaviour and weight loss: A meta-analysis and systematic review. *Obesity Reviews*, *20*(11), 1628–1641. https://doi.org/10.1111/obr.12916

Received November 19, 2019
Revision received July 4, 2025
Accepted July 10, 2025 ■