Attitudinal Effects of Stimulus Co-occurrence and Stimulus Relations: Range and Limits of Intentional Control

Bertram Gawronski & Skylar M. Brannon

University of Texas at Austin

Word Count (incl. abstract, main text, references, footnotes): 10,318

The reported research was supported by National Science Foundation Grant # 1649900 to Bertram Gawronski. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We thank Sarah Carr, Trey Delafosse, Sojung Lee, Katie Musemeche, and Yocelyn Rivera for their help in collecting the data.

Abstract

Research suggests that evaluations of an object can be simultaneously influenced by (1) the mere co-occurrence of the object with a pleasant or unpleasant stimulus (e.g., mere co-occurrence of object A and negative event B) and (2) the object's particular relation to the co-occurring stimulus (e.g., object A starts vs. stops negative event B). Using a multinomial modeling approach to disentangle the two kinds of influences on choice decisions, three experiments investigated whether learners can intentionally control the relative impact of stimulus co-occurrence and stimulus relations. An integrative analysis of the data from the three experiments (N = 1154) indicate that incentivized instructions to counteract effects of stimulus co-occurrence by focusing on stimulus relations increased the impact of stimulus relations without affecting the impact of stimulus co-occurrence. Implications for evaluative learning, intentional control, and public policy are discussed.

Keywords: associative learning; dual-process theory; evaluative conditioning; multinomial modeling; propositional learning

On March 17, 2020, the Food and Drug Administration of the United States issued a rule that requires tobacco companies to print graphic images depicting negative health consequences of smoking on cigarette packages (U.S. Food and Drug Administration, 2020). Following the success of similar policies in other countries (e.g., Australia, Canada), the rule resonates with the idea that repeatedly pairing cigarettes with unpleasant images might engender negative reactions towards cigarettes even when people reject the meaning of the graphic message (e.g., when people dismiss the proposition that smoking causes cancer; see Noar, Francis, Bridges, Sontag, Ribisil, & Brewer, 2016). Now, imagine a hypothetical proposal that aims to promote the use of sunscreen to prevent skin cancer by printing graphic images of skin cancer on sunscreen packages. Intuitively, one might be skeptical about the effectiveness of such a health campaign. Consistent with this intuition, research suggests that repeated co-occurrence of an object with a pleasant or unpleasant stimulus can lead to an evaluative response to the object that is congruent with the valence of the co-occurring stimulus even when the object is known to have a contrastive relation to the co-occurring stimulus (e.g., Heycke & Gawronski, 2020; Hu, Gawronski, & Balas, 2017; Kukken, Hütter, & Holland, 2020; Moran & Bar-Anan, 2013). For example, repeated pairings of sunscreen with unpleasant images of skin cancer (*stimulus co-occurrence*) may lead to negative responses toward sunscreen even when people understand and accept the intended message that sunscreen protects against skin cancer (*stimulus relation*).

The aim of the current research was to investigate the extent to which effects of stimulus co-occurrence and stimulus relations can be intentionally controlled. Using a formal modeling approach to disentangle the two kinds of influences on choice decisions, we were particularly interested in whether enhanced motivation to counteract effects of stimulus co-occurrence by focusing on stimulus relations is effective in reducing mere co-occurrence effects. Using the

opening example, does enhanced motivation to process the contrastive relation between sunscreen and skin cancer in the message *sunscreen prevents skin cancer* reduce negative responses to sunscreen resulting from the co-occurrence of *sunscreen* and *skin cancer* in the message?

Stimulus Co-occurrence and Stimulus Relations

Evidence for simultaneous effects of stimulus co-occurrence and stimulus relations comes from several studies using a task-dissociation approach. The central finding in this line of work is that implicit measures (e.g., evaluative priming, implicit association test) tend to reflect effects of stimulus co-occurrence, whereas explicit measures (e.g., evaluative rating scales) tend to reflect effects of stimulus relations (for an overview of implicit measures, see Gawronski & De Houwer, 2014). For example, in a study by Moran and Bar-Anan (2013), participants were repeatedly presented with sequences of images and sounds. Each sequence started with an image of one alien creature, followed by either a pleasant or an unpleasant sound (i.e., pleasant melody or unpleasant scream), followed by an image of a different alien creature. Participants were told that, depending on their position in the sequence, some aliens would start the following sound whereas other aliens would stop the preceding sound. Afterwards, evaluative responses to the alien creatures were measured with an explicit and an implicit measure. Whereas responses on the explicit measure reflected the particular relation of the aliens to the sounds, responses on the implicit measure reflected the mere co-occurrence of aliens and sounds regardless of their relation. Specifically, on the explicit measure, participants showed more favorable judgments of aliens that started pleasant sounds compared with aliens that stopped pleasant sounds. Conversely, participants showed less favorable judgments of aliens that started unpleasant sounds compared with aliens that stopped unpleasant sounds. In contrast, on the implicit

measure, participants showed more favorable responses to aliens that co-occurred with pleasant sounds compared with aliens that co-occurred with unpleasant sounds, regardless of whether the aliens started or stopped the sounds.

Similar findings were obtained by Hu et al. (2017, Experiments 1 and 2), who presented participants with image pairs involving pharmaceutical products and positive or negative health conditions. Participants were told that the pharmaceutical products either cause or prevent the depicted health conditions. Afterwards, evaluative responses to the pharmaceutical products were measured with an explicit and an implicit measure. Consistent with Moran and Bar-Anan's (2013) results, Hu et al. found that responses on the explicit measure reflected the relation between the pharmaceutical products and the depicted health conditions. In contrast, responses on the implicit measure reflected the mere co-occurrence of the products with the depicted health conditions regardless of their relation. Specifically, on the explicit measure, participants showed more favorable judgments of products that caused positive health conditions compared with products that prevented positive health conditions. Conversely, participants showed less favorable judgments of products that caused negative health conditions compared with products that prevented negative health conditions. In contrast, on the implicit measure, participants showed more favorable responses to products that co-occurred with positive health conditions than products that co-occurred with negative health conditions, regardless of whether the products caused or prevented the health conditions.

Although the findings by Moran and Bar-Anan (2013) and Hu et al. (2017) suggest a dissociation in the effects of stimulus co-occurrence and stimulus relations on implicit and explicit measures, the available evidence for unqualified co-occurrence effects on implicit measures is rather mixed (see Kurdi & Dunham, in press). Whereas some studies found mere co-

occurrence effects on implicit measures that remained unqualified by relational information (e.g., Hu et al., 2017, Experiments 1 and 2, Moran & Bar-Anan, 2013), other studies found attenuated co-occurrence effects when the co-occurring stimuli had a contrastive relation (e.g., Zanon, De Houwer, & Gast, 2012; Zanon, De Houwer, Gast, & Smith, 2014). Yet, other studies found a reversal of mere co-occurrence effects in cases involving contrastive relations (e.g., Gawronski, Walther & Blank, 2005; Hu et al., 2017, Experiment 3). An illustrative example of these inconsistencies is the work by Hu et al. (2017), who found differential effects of stimulus co-occurrence and stimulus relations on implicit and explicit measures only when the relational information was provided before the impression formation task and this information was consistent for all of the presented target stimuli (Experiments 1 and 2). However, when relational information was provided during the impression task and the specific relations varied on a trial-by-trial basis, both implicit and explicit measures were influenced by stimulus relations without showing any effect of stimulus co-occurrence (Experiment 3). Together with concerns about ambiguities in the theoretical meaning of dissociations between implicit and explicit measures (Corneille & Mertens, in press) and the conceptual distinction between implicit and explicit measures more broadly (Corneille & Hütter, 2020), these inconsistent findings raise significant questions about the suitability of a task-dissociation approach to disentangle effects of stimulus co-occurrence and stimulus relations (see Bading, Stahl, & Rothermund, 2020; Corneille & Stahl, 2019; Green, Luck, Gawronski, & Lipp, in press).

In line with these concerns, studies that used a multinomial modeling approach (see Hütter & Klauer, 2016) to disentangle effects of stimulus co-occurrence and stimulus relations have obtained more consistent evidence (Heycke & Gawronski, 2020; Kukken et al., 2020). Different from the comparison of responses across measures in the task-dissociation approach, a

central feature of the multinomial modeling approach is that it allows researchers to quantify independent contributions of stimulus co-occurrence and stimulus relations to overt responses on a single task. The two kinds of effects are captured by separate parameters quantifying the probabilities that responses reflect (1) a response pattern consistent with the observed stimulus relations and (2) a response pattern consistent with the observed stimulus co-occurrences.

For example, using a variant of Moran and Bar-Anan's (2013) learning paradigm, Kukken et al. (2020) found that participants' responses to the alien creatures were simultaneously influenced by both (1) their mere co-occurrence with a pleasant or unpleasant sound and (2) their particular relation to the co-occurring sound (i.e., whether they started or stopped the sound). Similarly, using a variant of Hu et al.'s (2017) learning paradigm, Heycke and Gawronski (2020) found that participants' responses to the pharmaceutical products were simultaneously influenced by both (1) their mere co-occurrence with a pleasant or unpleasant health condition and (2) their particular relation to the co-occurring health condition (i.e., whether they caused or prevented the health condition). Interestingly, Heycke and Gawronski obtained reliable effects of stimulus co-occurrence despite using a procedural setup that failed to produce mere co-occurrence effects on implicit measures in Hu et al.'s research (Experiment 3). Although studies using a multinomial modeling approach have identified several contextual conditions that moderate the relative impact of stimulus co-occurrence and stimulus relations (Heycke & Gawronski, 2020; Kukken et al., 2020), the obtained results support the idea that stimulus co-occurrence and stimulus relations jointly influence evaluative responses.

Theoretical Explanations

A common explanation of the distinct effects of stimulus co-occurrence and stimulus relations is that they are the products of two functionally distinct learning mechanisms. For

example, according to the associative-propositional evaluation (APE) model (Gawronski & Bodenhausen, 2006, 2011, 2018), mere co-occurrence effects are the product of an associative learning mechanism involving the automatic formation of mental associations between co-occurring stimuli. In contrast, effects of stimulus relations are claimed to be the product of a propositional learning mechanism involving the non-automatic generation and truth assessment of mental propositions about the relation between co-occurring stimuli. Based on the hypothesis that effects of stimulus co-occurrence and stimulus relations are mediated by two distinct learning mechanisms, such accounts have been described as *dual-process learning accounts*.

An alternative explanation is offered by theories that interpret all learning effects as outcomes of a single propositional mechanism involving the non-automatic generation and truth assessment of mental propositions about stimulus relations (e.g., De Houwer, 2009, 2018; De Houwer, Van Dessel, & Moran, 2020). According to these theories, distinct effects of stimulus co-occurrence and stimulus relations result from processes during the retrieval of stored propositional information rather than two functionally distinct learning mechanisms. For example, based on the assumptions of the Integrated Propositional Model (IPM; De Houwer, 2018), mere co-occurrence effects can be expected to occur despite the successful learning of contrastive information when the retrieval of stored propositions about stimulus relations is incomplete (e.g., retrieval of *A is related to B* rather than *A stops B*; see Van Dessel, Gawronski, & De Houwer, 2019). Based on the hypothesis that effects of stimulus co-occurrence and stimulus relations can arise from a single propositional learning mechanism, such accounts have been described as *single-process learning accounts*.

Intentional Control

Some researchers suggested that is impossible to empirically distinguish between dual-process and single-process accounts, because any finding that conflicts with the predictions derived from a given theory may be reconciled with that theory by means of post-hoc assumptions (De Houwer et al., 2020). Nevertheless, tests of competing predictions derived from the core assumptions of dual-process versus single-process accounts can be valuable, because (1) such tests provide novel empirical insights and (2) post-hoc assumptions proposed to explain an unpredicted finding can generate new empirical research to test these ad hoc assumptions (Gawronski & Bodenhausen, 2015a). Expanding on these ideas, the current research tested competing predictions derived from dual-process and single-process accounts about whether the relative impact of stimulus co-occurrence and stimulus relations can be intentionally controlled. Using a multinomial modeling approach to disentangle effects of stimulus co-occurrence and stimulus relations, we were especially interested in how enhanced motivation to counteract effects of stimulus co-occurrence by focusing on stimulus relations influences the relative impact of the two kinds of information. Based on the conflicting assumptions of the APE model and the IPM regarding the existence of two functionally distinct learning mechanisms, the current research focused on the impact of intentional control during learning (i.e., formation of evaluative representation), while minimizing opportunities for intentional control during judgment (i.e., expression of evaluative representation).

From the perspective of the APE model, enhanced motivation to counteract effects of stimulus co-occurrence by focusing on stimulus relations should strengthen the impact of relational information via enhanced propositional processing of the to-be-learned relations. However, it should have little impact on the effect of stimulus co-occurrence, which is claimed

to result from the automatic formation of mental associations between co-occurring stimuli. Because the two learning processes are assumed to be independent, a stronger effect of stimulus relations resulting from enhanced propositional processing should have little impact on the automatic associative effect of stimulus co-occurrence. Based on these assumptions, enhanced motivation to counteract effects of stimulus co-occurrence by focusing on stimulus relations should increase the impact of stimulus relations without affecting the impact of stimulus co-occurrence.

A different set of predictions can be derived from the core assumptions of the IPM. According to the IPM, effects of stimulus co-occurrence are due to incomplete retrieval of stored relational information (e.g., retrieval of *A is related to B* rather than *A stops B*) rather than two functionally distinct learning mechanisms. Yet, contextual conditions during learning may influence the retrieval of stored relational information by influencing the storage of relational information in long-term memory. Specifically, enhanced processing of stimulus relations during learning should support the storage of relational information in long-term memory, and improved storage of relational information in long-term memory should reduce the likelihood of incomplete retrieval of the stored relational information. Together these assumptions imply that enhanced motivation to counteract effects of stimulus co-occurrence by focusing on stimulus relations should increase the impact of stimulus relations and reduce the impact of stimulus co-occurrence.

The Current Research

To test the competing predictions derived from the APE model and the IPM, the current research used Heycke and Gawronski's (2020) RCB model to disentangle effects of stimulus co-occurrence and stimulus relations (see also Kukken et al., 2020). Following the procedure of Hu

et al.'s (2017, Experiment 3) learning paradigm, participants were presented with pairings of pharmaceutical products (i.e., conditioned stimuli, CSs) and images of positive or negative health conditions (i.e., unconditioned stimuli, USs). For half of the pairings, participants received information that the pharmaceutical product causes the depicted health condition. For the remaining half, participants received information that the pharmaceutical product prevents the depicted health condition. Participants' task was to form an impression of the pharmaceutical products based on the presented information. Afterwards, participants were presented with the pharmaceutical products one-by-one and asked to indicate whether or not they would choose the product (*yes* vs. *no*).

Applied to Hu et al.'s (2017) learning paradigm, the RCB model captures patterns of evaluative responses to four kinds of stimuli: (1) pharmaceutical products that cause positive health outcomes, (2) pharmaceutical products that cause negative health outcomes, (3) pharmaceutical products that counteract positive health outcomes, and (4) pharmaceutical products that counteract negative health outcomes (see Figure 1). Based on the observed responses to the four kinds of stimuli, the model provides numerical estimates for the probabilities that (1) responses to the pharmaceutical products are driven by their relation to the depicted health outcomes (labeled *R*), (2) responses to the pharmaceutical products are driven by their mere co-occurrence with the depicted health outcomes (labeled *C*), and (3) responses to the pharmaceutical products reflect a general positivity or negativity bias regardless of their relation and co-occurrence with particular health outcomes (labeled *B*).

To investigate the impact of intentional control on the effects of stimulus co-occurrence (captured by the RCB model's *C* parameter) and stimulus relations (captured by the RCB model's *R* parameter), half of the participants were instructed to avoid being influenced by the

mere co-occurrence of the pharmaceutical products and the depicted health conditions. To avoid such influences, participants were instructed to focus on the causal relations of the pharmaceutical products to the depicted health conditions (i.e., whether a product causes or prevents the co-occurring health condition). As an incentive for their efforts, participants were told that we would give a $100 bonus to the participant with the best performance on the task. The remaining half completed the learning task without control-instructions. To isolate effects of intentional control during learning, we presented the control-instructions before the learning task and minimized opportunities for intentional control in the judgment task by assessing choice responses under time pressure.[1]

To investigate the impact of intentional control on the effects of stimulus co-occurrence and stimulus relations, we conducted three experiments. For each study, we aimed to recruit 400 participants. For our between-subjects manipulation of intentional control, a sample of 400 participants provides a power of 80% in detecting a small effect of $d = 0.28$ in a traditional *t*-test for independent means (two-tailed).[2] To increase statistical power for the detection of smaller effects, we also conducted an integrative data analysis (IDA; see Curran & Hussong, 2009) using the combined sample from all three studies. The combined sample after exclusions ($N = 1154$) provides a power of 95% in detecting a small effect of $d = 0.21$ in a traditional *t*-test for independent means (two-tailed). Because the critical difference between the APE model and the IPM involves the absence versus presence of a significant effect of control-instructions on the *C*

[1] Both the APE model and the IPM predict that intentional control during judgment should increase effects of stimulus relations and decrease effects of stimulus co-occurrence (see Heycke & Gawronski, 2020). To the extent that opportunities for intentional control during judgment increase as a function of available time (see Moors, 2016), the two theories predict different outcomes only for conditions of time pressure, but not for conditions of unlimited time.

[2] Because power analyses within multinomial modeling require simulations with expected population values for the three parameters and any specific expectations in this regard would be arbitrary, we made our a priori sample-size decision in a heuristic fashion based on simple comparisons of mean values using *t*-tests.

parameter, we report the results of the high-powered IDA in the main article and the results of the three individual experiments in the Supplemental Materials.[3] The data for each study were collected in one shot without intermittent statistical analyses. We report all measures, all conditions, and all data exclusions. The materials, raw data, and analysis files for all studies are publicly available at https://osf.io/cuaz6/?view_only=29d15e36b27443738ee70bf08fcbc0ff. The protocol of the three experiments was approved by the Institutional Review Board of [blinded for review] under protocol #2016-11-0092.

Methods

Participants and Design

All three experiments included the same 2 (US Valence: positive vs. negative) × 2 (CS-US Relation: causes vs. prevents) × 2 (Task Instructions: standard vs. control) mixed design with the first two variables being manipulated within-subjects and the last one being manipulated between-subjects. Experiment 1 was conducted as a lab study; Experiments 2 and 3 were conducted as online studies. The combined sample for the IDA included 1154 participants (633 women, 519 men, 2 other), with 582 participants in the standard-instructions condition and 572 participants in the control-instructions condition.

For Experiment 1, we recruited 413 psychology undergraduates for a one-hour battery entitled "First Impressions" that included the current study and two unrelated studies.[4] The current study was always completed as the first one in the battery. Participants received credit for

[3] The results of the three individual experiments converge with the results of the IDA, the only exceptions being that (1) the RCB model did not fit the data in Experiment 1 and (2) the effect of intentional control on the *R* parameter was not significant in Experiment 2.

[4] Due to excessive sign-ups at the end of the academic term, the sample size was slightly larger than the desired sample size of 400 participants.

a research participation requirement. Due to experimenter error, data from one participant were lost, leaving us with valid data from 412 participants (274 women, 138 men).

Participants for Experiment 2 were recruited via Amazon's MTurk for a study entitled "How Do We Form Impressions of Novel Objects?". Eligibility for participation was limited to MTurk workers in the United States with a HIT approval rate of at least 95% who did not participate in prior studies from our lab using the same paradigm. Participants received compensation of $2.00 for completing the study. Of the 429 MTurk workers who began the study, 403 completed all measures. Four cases with duplicate subject codes (presumably due to multiple completions by the same participant) were excluded from analyses. Twelve participants failed to pass an instructional attention check (see below), 4 participants reported not paying attention to the images or not taking their responses seriously (see below), and 3 participants had invalid responses on more than 50% of the trials in the choice task. Following the exclusion criteria by Heycke and Gawronski (2020), data from these participants were excluded from the analyses, leaving us with valid data from 380 participants (172 women, 206 men, 2 other).

Participants for Experiment 3 were recruited via Amazon's MTurk following the procedures and eligibility criteria in Experiment 2. Participants received compensation of $2.00 for completing the study. Of the 424 MTurk workers who began the study, 403 completed all measures. Two cases with duplicate subject codes (presumably due to multiple completions by the same participant) were excluded from analyses. Twenty-four participants failed to pass an instructional attention check (see below), 10 participants reported not paying attention to the images or not taking their responses seriously (see below), and 5 participants had invalid responses on more than 50% of the trials in the choice task. Following the exclusion criteria by

Heycke and Gawronski (2020), data from these participants were excluded from the analyses, leaving us with valid data from 362 participants (187 women, 175 men).

Learning Task

Participants in all three experiments completed the same learning task, which was directly adapted from Heycke and Gawronski (2020). The task included information about whether pharmaceutical products cause or prevent either healthy or unhealthy physical conditions. The stimuli in the task included 12 images of hypothetical pharmaceutical products, 6 images of healthy physical conditions (e.g., voluminous hair), and 6 images of unhealthy physical conditions (e.g., tooth decay). On each trial of the task, an image of a pharmaceutical product (CS) was presented on the left and an image of a healthy or unhealthy physical condition (US) on the right, with one of the two qualifiers *causes* or *prevents* being presented in the center of the screen between the two images. Each stimulus combination was presented for 3000 ms with an inter-trial interval of 1000 ms. Three CSs were presented with a positive US and the qualifier *causes*; three CSs were presented with a negative US and the qualifier *causes*; three CSs were presented with a positive US and the qualifier *prevents*; and three CSs were presented with a negative US and the qualifier *prevents*. The use of a given CS for pairings with positive versus negative USs and the qualifiers *causes* versus *prevents* was counterbalanced by means of a Latin square. The learning phase consisted of 4 blocks with self-paced breaks between blocks. Within each block, each CS-US-qualifier combination was presented twice, summing up to 8 presentations of each stimulus combination over the four blocks. For each participant, a given CS was always presented together with the same US. With 12 unique CS-US-qualifier combinations and 8 presentations of each CS-US-qualifier combination, the learning task included a total of 96 trials.

Task-Instructions Manipulation

To investigate effects of intentional control, participants in the three experiments were randomly assigned to either a *standard-instructions* condition or a *control-instructions* condition. Participants in both conditions received the same basic instructions before the learning task:

> *The next part of this study is concerned with how people process information about consumer products. For this purpose, you will be presented with images of pharmaceutical products and visual information about their effects. As you know, many pharmaceutical products have positive effects, but some products also have negative side-effects. For each product you will see whether this product causes or prevents a health outcome. Your task is to think of the image pairs, such that the pharmaceutical product CAUSES or PREVENTS what is displayed in the other photograph. For example, if a product is paired with a positive image and it says 'causes', you should think of the product in terms of it causing the positive outcome displayed in the image. Conversely, if a product is paired with a negative image and it says 'causes', you should think of the product in terms of it causing the negative outcome displayed in the image. If a product is paired with a positive image and it says 'prevents', you should think of the product in terms of it preventing the positive outcome displayed in the image. Conversely, if a product is paired with a negative image and it says 'prevents', you should think of the product in terms of it preventing the negative outcome displayed in the image. Again, please think of the image pairs in the relation mentioned on the screen (causes or prevents). The task will take approximately 5 minutes.*

In Experiment 1, participants in the control-instructions condition received the following information in addition to the basic instructions (see Gawronski, Balas, & Creighton, 2014):

IMPORTANT!!! Previous research suggests that repeated pairings of a pharmaceutical product with pleasant or unpleasant images can influence people's responses to the pharmaceutical products regardless of their causal relation. Specifically, it has been shown that responses to a pharmaceutical product become more positive when the product is repeatedly paired with a pleasant image, regardless of whether the product causes or prevents the pleasant health condition displayed in the image. Conversely, responses to a pharmaceutical product become more negative when the product is repeatedly paired with a negative image, regardless of whether the product causes or prevents the unpleasant health condition displayed in the image. In the current study, we are interested in whether such "evaluative conditioning" effects can be eliminated by people's intentional efforts to form impressions in line with the causal relation between the pharmaceutical products and the depicted health conditions. That is, can people avoid being influenced by repeated pairings of a pharmaceutical product with pleasant or unpleasant images by focusing on its causal relation to the depicted health conditions (i.e., whether the product causes or prevents the depicted health condition)? As an incentive for your efforts, we will give a $100 Amazon gift card to the participant who shows the best performance on this task. If you want your data to be considered for our performance-based incentive, please contact the experimenter to obtain a personal code. You will be asked to include your personal code on the next screen, so that we can contact the winner of the $100 Amazon gift card without having to record any identifying information from our participants. After completion of the study, we will send a mass email with the personal code of the winner to all participants in this study. The winner will be asked to identify him- or herself by showing us the stamped slip with the personal

code. Please contact the experimenter now to obtain a stamped slip with your personal code and include the code in the text box below. If you do not want to be considered for the $100 Amazon gift card, please type "no" in the text box below and click "Continue".

After participants in the control-instructions condition obtained their personal code and typed it into the text box, they received a short reminder before they were asked to start the learning task:

Again, please avoid being influenced by the repeated pairings of the pharmaceutical products with pleasant or unpleasant images by focusing on the causal relation between the products and the depicted health conditions (i.e., whether a product causes or prevents the depicted health condition). The task will take approximately 5 minutes.

To ensure anonymity, the program did not record the personal code participants entered in the text box and the winner of the $100 gift card was selected by means of a random procedure. For the sake of fairness, participants in the standard-instructions condition were included in the lottery for the gift card. Toward this end, participants in the standard-instructions condition received a personal code after the manipulation checks (see below) and asked to type their personal code in a text box on the computer screen. As with the control-instructions condition, the personal code was not recorded to ensure anonymity. The randomly selected winner of the gift card was announced after completion of the study via a mass email to all participants. The email included the personal code of the winner, asking the winner to contact a research assistant to schedule a time to pick up the gift card. The winner was required to provide their stamped slip with the personal code upon pick-up.[5]

[5] To prevent fraudulent claims of the gift card by means of fabricated slips with the winning code, the slips included a university stamp of the first author's lab.

In Experiments 2 and 3, the instructions were identical, the only difference being that, instead of promising a $100 Amazon gift card to the participant with the best performance, participants in the control-instructions condition were told that the participant with the best performance would receive a bonus payment of $100 to their MTurk account. Participants in the standard-instructions condition were told that we will have a lottery for a $100 bonus as a token of appreciation for their participation in the study. The winner of the $100 bonus was identified by means of a random procedure. The bonus was transferred to the winner's MTurk account after completion of the study. After the transfer, a mass email was sent to all participants that the funds had been transferred to the winner's account, identifying the winner with the last five digits of their MTurk worker ID.

Measures

Choice task. After the learning task, participants in all three experiments completed a speeded choice task in which they were asked to indicate whether they would choose a given product (see Heycke & Gawronski, 2020). On each trial of the task, a CS was shown in the center of the screen, and participants had 1000 ms to indicate whether or not they would choose the presented product. Participants were asked to press a left-hand key (*A*) if their answer was *no* and a right-hand key (*Numpad 5* in Experiment 1; *K* in Experiments 2 and 3) if their answer was *yes*. If participants did not respond within the 1000 ms response window, a short message was displayed in the center of the screen. In Experiment 1, participants were presented with the message *Please try to respond faster!* for 1000 ms. In Experiments 2 and 3, participants were presented with the message *Too slow* for 750 ms. Only valid responses within the 1000 ms response window were used in the analysis. Each trial started with a blank screen (presented for 500 ms in Experiment 1 and for 100 ms in Experiments 2 and 3), followed by a fixation cross

(presented for 500 ms in Experiment 1 and for 900 ms in Experiments 2 and 3). During the 1000 ms presentation of a given CS, labels for the two response options (*no* vs. *yes*) were displayed on the bottom-left side and the bottom-right side of the screen, with the question *Would you choose this product?* being displayed slightly below the CS. The choice task included three blocks, with each CS being presented once in each block, summing up to a total of 36 trials. The order of CSs within each block was randomized separately for each participant.

Manipulation checks. To test the effectiveness of the task-instructions manipulation, participants were asked to answer two questions after the choice task. The first item asked participants to rate their motivation to form impressions of the pharmaceutical products that are in line with the depicted causal relation to the health outcomes. The second item asked participants to rate their motivation to avoid being influenced by the mere pairings of the pharmaceutical products and the depicted health outcomes. Responses to both items were recorded with 7-point rating scales ranging from 1 (*not at all*) to 7 (*very much*).

Attention checks. Following the procedures by Heycke and Gawronski (2020), the two online studies (Experiments 2 and 3) included three measures to identify participants who did not pay sufficient attention. The first measure was a one-item instructional attention check (Oppenheimer, Meyvis, & Davidenko, 2009) with the following instructions:

> *Most modern theories of decision-making recognize the fact that decisions do not take place in a vacuum. Individual preferences and knowledge, along with situational variables can greatly impact the decision process. In order to facilitate our research on decision-making we are interested in knowing certain factors about you, the decision maker. Specifically, we are interested in whether you actually take the time to read the directions; if not, then some of our manipulations that rely on changes in the instructions*

will be ineffective. So, in order to demonstrate that you have read the instructions, please ignore the sports items below. Instead, simply continue on to the next page after the options. Thank you very much.

Below the instructions, participants were presented with the question *Which of these activities do you engage in regularly? (check all that apply)* and the response options: *Football, Soccer, Dancing, Watersports, Triathlon, Running, Volleyball*. By default, we excluded all participants from the analyses who, counter to the instructions, checked one or more of the response options on this item. In addition to the instructional attention check, participants were asked (1) if they paid attention to the images presented throughout the task and (2) if they took their responses in the study seriously (see Heycke & Gawronski, 2020). Participants were informed that their responses on these two items would not affect their compensation. By default, we excluded all participants from the analyses who reported that they did not pay attention to the images or did not take their responses seriously (see Aust, Diedenhofen, Ullrich, & Musch, 2013).

RCB Model

Because the mathematical underpinnings of the RCB model are explained in detail by Heycke and Gawronski (2020), we will only summarize the basic steps in analyzing data with the model. Based on the processing tree depicted in Figure 1, the RCB model provides four non-redundant mathematical equations to estimate numerical values for the three model parameters (*R*, *C*, *B*) based on the empirically observed probabilities of a *positive* versus *negative* response to the four types of stimuli (see Appendix in Heycke & Gawronski, 2020). These equations include the three model parameters as unknowns and the empirically observed probabilities of *positive* versus *negative* responses to the four types of stimuli as known numerical values. Using

maximum likelihood statistics, multinomial modeling generates parameter estimates for the three unknowns that minimize the difference between the empirically observed probabilities of positive versus negative responses to the four types of stimuli and the probabilities of positive versus negative responses predicted by the model equations using the generated parameter estimates. The adequacy of the model in describing the data can be evaluated by means of goodness-of-fit statistics, such that poor model fit would be reflected in a statistically significant deviation between the empirically observed probabilities in a given data set and the probabilities predicted by the model for this data set. Differences in parameter estimates across groups can be tested by enforcing equal estimates for a given parameter across groups. If setting a given parameter equal across groups leads to a significant reduction in model fit, it can be inferred that the parameter estimates for the two groups are significantly different. If setting a given parameter equal across groups does not lead to a significant reduction in model fit, the parameters for the two groups are not significantly different from each other. RCB model analyses were conducted with the free software multiTree v0.43 (Moshagen, 2010) and the template files provided by Heycke and Gawronski (2020) at https://osf.io/7ac4d/.

Results

Manipulation Checks

In line with the intended effect of the task-instruction manipulation, participants in the control-instructions condition reported a significantly stronger motivation to avoid being influenced by mere pairings than participants in the standard-instructions condition (*M*s = 5.22 vs. 3.78, respectively), $t(1152) = 13.43$, $p < .001$, $d = 0.79$. However, participants in the two conditions did not significantly differ in terms of their motivation to form impressions in line

with the depicted causal relations, which tended to be relatively high in both groups (*M*s = 5.47 vs. 5.42, respectively), $t(1152) = 0.65$, $p = .513$, $d = 0.04$.

Traditional Analysis

The choice data were aggregated by calculating the relative proportions of *yes* vs. *no* responses for each of the four categories of CSs within each of the two task-instructions conditions (see Table 1). Submitted to a 2 (US Valence) × 2 (CS-US Relation) × 2 (Task Instructions) mixed ANOVA, choice scores revealed a significant main effect of Task Instructions, $F(1, 1152) = 10.14$, $p = .001$, $\eta_G^2 = .009$, indicating that participants were more likely to choose the CSs in the control-instructions condition compared to the standard-instructions condition. There was also a significant main effect of US Valence, $F(1, 1152) = 75.16$, $p < .001$, $\eta_G^2 = .061$, indicating that participants were more likely to choose CSs paired with positive USs compared to CSs paired with negative USs. The main effect of US Valence was qualified by a significant two-way interaction between US Valence and CS-US Relation, $F(1, 1152) = 246.85$, $p < .001$, $\eta_G^2 = .176$. Post-hoc tests showed that, when the CSs were described as causing the USs, CSs paired with positive USs were chosen more frequently than CSs paired with negative USs, $t(1153) = 16.27$, $p < .001$, $d = 0.479$. Conversely, when the CSs were described as preventing the USs, CSs paired with positive USs were chosen less frequently than CSs paired with negative USs, $t(1153) = -6.01$, $p < .001$, $d = 0.177$. Moreover, when the CSs were paired with positive USs, CSs that were described as causing the USs were chosen more frequently than CSs that were described as preventing the USs, $t(1153) = 13.48$, $p < .001$, $d = 0.397$. Conversely, when the CSs were paired with negative USs, CSs that were described as causing the USs were chosen less frequently than CSs that were described as preventing the USs, $t(1153) = -12.76$, $p < .001$, $d = 0.376$. The three-way interaction between US Valence, CS-US

Relation, and Task Instructions was marginal, $F(1, 1152) = 3.77$, $p = .052$, $\eta_G^2 = .003$, indicating that the two-way interaction between US Valence and CS-US Relation tended to be more pronounced in the control-instructions condition, $F(1, 571) = 136.92$, $p < .001$, $\eta_G^2 = .193$, compared to the standard-instructions condition, $F(1, 581) = 109.43$, $p < .001$, $\eta_G^2 = .158$.

RCB Model

The RCB model was fit to the data from all three experiments with the three model parameters varying freely across task-instructions conditions. Despite the large sample size (N = 1154) and the high statistical power in detecting even minor deviations between predicted and observed responses, the model fit the data well, $G^2(2) = 3.90$, $p = .142$, $w = .010$. This model was used as a baseline for tests whether the three model parameters are significantly different across task-instructions conditions (see Table 2). Analyses revealed a significant effect of Task Instructions on the B parameter, $\Delta G^2(1) = 35.11$, $p < .001$, $w = .030$, indicating that participants in the standard-instructions condition had a stronger response bias to reject the products than participants in the control-instructions condition. More important for the current question, a significant effect of Task Instructions on the R parameter indicated that relational information had a greater impact on participants' choices in the control-instructions condition compared to the standard-instructions condition, $\Delta G^2(1) = 10.82$, $p = .001$, $w = .017$. There was no significant effect of Task Instructions on the C parameter, $\Delta G^2(1) = 0.11$, $p = .740$, $w = .002$. The effect of Task Instructions was significantly different for the R and C parameters, $\Delta G^2(1) = 10.93$, $p < .001$, $w = .017$.

Discussion

The main goal of the current research was to investigate the extent to which attitudinal effects of stimulus co-occurrences and stimulus relations can be intentionally controlled. Overall,

we found that instructions to counteract effects of stimulus co-occurrence by focusing on stimulus relations enhanced the impact of stimulus relations while being ineffective in reducing the impact of stimulus co-occurrences. These findings are consistent with predictions derived from the APE model (Gawronski & Bodenhausen, 2006, 2011, 2018), which suggests that effects of stimulus co-occurrence are driven by an associative learning mechanism and effects of stimulus relations are driven by a propositional learning mechanism. In contrast, the findings are inconsistent with predictions derived from the IPM (De Houwer, 2018), which postulates a single propositional learning process whose behavioral outcomes depend on the (in)complete retrieval of stored propositional information. Whereas the APE model predicts that instructions to counteract effects of stimulus co-occurrences by focusing on stimulus relations should enhance effects of stimulus relations without reducing effects of stimulus co-occurrences, the IPM suggests that such instructions should enhance effects of stimulus relations and reduce effects of stimulus co-occurrences.

Although the current findings conflict with the predictions derived from the IPM, it is worth noting that the model is sufficiently flexible to be reconciled with a wide range of conflicting findings in a post-hoc fashion (see De Houwer et al., 2020). One potential way to reconcile the IPM with the current findings is to propose that (1) people generate and store two propositions for the same event, one capturing relational information (e.g., *X prevents something negative*) and one capturing co-occurrence information (e.g., *X co-occurs with something negative*), and (2) propositions capturing co-occurrence information are generated and stored automatically. Although these post-hoc assumptions reconcile the IPM with the current findings, it is worth noting that they make the theory empirically indistinguishable from theories that propose two functionally distinct learning mechanisms, rendering the debate a matter of

terminological preference rather than empirical evidence. While dual-process learning theories explain mere co-occurrence effects in terms of automatic formation of associations between co-occurrence stimuli, the post-hoc explanation provided by IPM would explain mere co-occurrence effects in terms of automatic processing of co-occurrence propositions.

These considerations echo concerns that, in the absence of precise hypotheses about (1) the contents of propositions generated during learning and (2) the conditions that influence their storage and retrieval, single-process propositional theories are too flexible to prohibit specific empirical outcomes (see Kurdi & Dunham, in press). Nevertheless, single-process propositional theories have the potential to generate novel insights by inspiring empirical studies that seem unlikely to be conducted without their theoretical guidance (see De Houwer et al., 2020). Indeed, we probably would not have conducted the current studies if dual-process theories such as the APE model had not been challenged by the findings of studies inspired by single-process propositional theories (e.g., Hu et al., 2017; Peters & Gawronski, 2011; for a review, see Corneille & Stahl, 2019). It is also worth noting that, although the current findings support predictions derived from the APE model and conflict with predictions derived from the IPM, both theories have difficulties in explaining the findings of other studies that have used a multinomial modeling approach to disentangle effects of stimulus co-occurrence and stimulus relations (Heycke & Gawronski, 2020). Thus, although the APE model has superior explanatory power for the current findings, both the APE model and the IPM are facing significant empirical challenges that require non-trivial theoretical revisions.

Implications for Intentional Control

By investigating the impact of intentional control on the effects of stimulus co-occurrence and stimulus relations, the current work expands on earlier research on the controllability of

mere co-occurrence effects in evaluative conditioning (EC). Different from the contrasting of stimulus co-occurrence and stimulus relations in the current studies, earlier research investigated whether repeated pairings of a CS with a positive or negative US influence evaluative responses to the CS even when participants are instructed to avoid being influenced by the pairings. Using a task-dissociation approach, Gawronski et al. (2014) found that control-instructions moderated EC effects on an explicit measure without affecting EC effects on an implicit measure. Similar findings were obtained by Hütter and Sweldens (2018) who used a multinomial modeling approach to disentangle controlled and uncontrolled influences of CS-US pairings on evaluative judgments (for related findings, see Balas & Gawronski, 2012; Corneille, Mierop, Stahl, & Hütter, 2019; Gawronski, Mitchell, & Balas, 2015). The current findings provide further insights into the limits of intentional control, showing that intentional control can enhance effects of stimulus relations without reducing effects of stimulus co-occurrences.

As a caveat, it is important to note that the current findings provide evidence for *uncontrolled* effects of stimulus co-occurrences, but this evidence does not necessarily imply that effects stimulus co-occurrences are *uncontrollable*. After all, it is possible that participants in the current studies adopted a suboptimal strategy to avoid effects of stimulus co-occurrences (e.g., strategically reduced response bias during judgment, as reflected in a significant effect on the *B* parameter) and that a different control strategy might have been more effective. However, previous evidence regarding the controllability of mere co-occurrence effects in EC gives reasons to remain skeptical about this possibility. Using a task-dissociation approach, Gawronski et al. (2015) investigated the effectiveness of three emotion-focused control strategies in reducing EC effects on explicit and implicit measures: (1) suppression of emotional reactions to the US, (2) reappraisal of US valence, and (3) facial blocking of emotional expressions.

Although all three strategies reduced EC effects on explicit measures via impaired memory for CS-US pairings, none of them was effective in reducing EC effects on an implicit measure. Future research may provide further insights into the effectiveness of different control strategies by investigating their impact on the effects of stimulus co-occurrence and stimulus relations using a multinomial modeling approach.

Implications for Public Policy

The current findings have important implications not only for the theoretical debate between dual-process and single-process learning theories, but also for public policy and the implementation of federal trade laws. The U.S. Federal Trade Commission explicitly bans advertisement techniques that violate the ethical principle of consumer autonomy, which states that consumers should have the ability to determine their own destiny (Nebenzahl & Jaffe, 1998). To the extent that (1) mere co-occurrences of stimuli can influence judgments and decisions in a manner that conflicts with the encoded meaning of stimulus relations, and (2) such co-occurrence effects cannot be intentionally controlled, advertisements involving effects of mere co-occurrence would be in violation with the principle of consumer autonomy. In line with this concern, the current findings suggest that top-down processes during encoding (e.g., enhanced processing of stimulus relations) may be ineffective in reducing effects of stimulus co-occurrences. However, it is worth noting that, in order to isolate effects of intentional control during learning and minimize effects of intentional control during judgment, participants in the current studies had to make their decisions under time pressure. Thus, participants might have been more successful in controlling effects of stimulus co-occurrences if they had been given more time during judgment. Although previous findings suggest that more time during judgment increases (rather than decreases) effects of stimulus co-occurrence (Heycke & Gawronski, 2020),

future research is needed to determine the cognitive requirements for effective intentional control of co-occurrence effects during learning and judgment (see Footnote 1).

Potential Objections

Although the current findings are consistent with predictions derived from the APE model and inconsistent with predictions derived from the IPM, it is seems appropriate to address some potential objections to our conclusions. First, our main conclusion is based on a null effect of intentional control on the *C* parameter, which could be due to multiple factors other than automatic association formation. One such factor is insufficient statistical power. In the current research, we aimed to address this concern by conducting three independent replications with large sample sizes and by reporting the results of an IDA that used the combined sample from all three studies ($N = 1154$). Yet, even the high-powered IDA did not obtain a significant effect of control instructions on the *C* parameter. Although it is possible that the impact of control instructions on the effect of stimulus co-occurrence is too small to be detected with the combined sample, it seems debatable if such small effects impose meaningful constraints on theories about underlying mental processes.

A second potential concern is that the four cases in the manipulation of US Valence and CS-US relations are not comparable, because some cells seem more difficult to process than others. Although there may be mental models for cases in which pharmaceutical products cause positive outcomes, counteract negative outcomes, and cause negative outcomes, the case of pharmaceutical products counteracting positive outcomes may seem unusual and thus more difficult to process. There are two reasons why this objection does not qualify the current conclusions. First, if participants have difficulties in mentally representing the identified case, a basic assumption of the RCB model would be violated, which should undermine the fit of the

model in the describing the data. Yet, counter to this concern, the RCB model fit the data well despite the high statistical power in detecting even minor deviations between predicted and observed responses in the IDA. Second, the presumed asymmetry should negatively affect the reliability of the *R* parameter, but it has no implications for the reliability of the *C* parameter, the latter of which depends exclusively on the valence of the USs. Hence, potential asymmetries between the four cases should reduce the likelihood of detecting effects on the *R* parameter, but not the *C* parameter. Yet, counter to this concern, the manipulation of intentional control showed significant effects on the *R* parameter, but not the *C* parameter.

A third concern is whether it is actually possible to ignore the mere co-occurrence of two stimuli in processing their relation. Numerous studies suggest that false information continues to influence judgments and decisions after being debunked (see Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012), and that enhanced elaboration increases the effect of the debunking message without reducing the impact of the debunked information (see Chan, Jones, Jamieson, & Albarracín, 2017). The current findings show a similar pattern, in that (1) "false" co-occurrence information influences choices despite "true" relational information and (2) greater elaboration of "true" relational information increases the impact of relational information without reducing the impact of "false" co-occurrence information. From this perspective, the current findings could be regarded as another demonstration of the known robustness of continued-influence effects. However, such a categorization provides only a different description of the observed results without offering a theoretical explanation in terms of underlying mental processes (De Houwer, 2011; Gawronski & Bodenhausen, 2015). While the APE model offers a mental process account that predicts the current findings in an a priori fashion, a mental process explanation in

terms of the IPM requires post-hoc assumptions that make the theory indistinguishable from a dual-process account (see above).

A fourth concern is that one of the two manipulation checks consistently failed in the individual studies as well as the IDA, raising questions about the validity of our instruction manipulation. In evaluating this concern, we deem it important to consider (1) the content of the control-instructions, (2) the content of the two manipulation checks, and (3) the nature of the observed asymmetry. The control-instructions asked participants to avoid being influenced by the mere co-occurrence of the pharmaceutical products and the depicted health outcomes. Toward this end, they were instructed to focus on the causal relation between the products and the depicted health conditions. The asymmetry in the manipulation checks suggests that the instructions manipulation effectively influenced participants' motivation to avoid being influenced by the mere co-occurrence of the pharmaceutical products and the depicted health outcomes. Yet, participants were highly motivated to form impressions in line with the depicted causal relations regardless of control-instructions. Interestingly, the increased motivation to counteract effects of mere co-occurrence increased participants' success in forming impressions in line with the depicted causal relations, but it did not reduce mere co-occurrence effects. In other words, our manipulation effectively enhanced the goal of counteracting effects of stimulus co-occurrence, but participants did not succeed in accomplishing this particular goal. Nevertheless, it did increase their success in accomplishing a salient goal that was equal across the two conditions: the goal of forming impressions in line with the depicted stimulus relations. Does this pattern question the validity of our experimental manipulation? We would argue that the answer to this question is *no, it does not*. It simply suggests a more complex relation between

the operation of task-relevant goals and goal achievement, but it does not question the validity of our manipulation in influencing the goal to avoid being influenced by mere co-occurrence.

A final concern is that our manipulation of intentional control included multiple components, which makes it difficult to identify which of these components was essential for the obtained results. First, the control-instructions were much longer compared the standard-instructions. Second, participants were informed about mere co-occurrence effects in the control-instructions condition, but not in the standard-instructions condition. Third, participants in the control-instructions condition were asked to devote extra efforts to processing relational information. Fourth, normatively accurate performance was incentivized in the control-instructions condition, but not in the standard-instructions condition. We believe that some of these components were more influential than others. Although it seems possible that the longer text in the control-instructions condition diluted the impact of the shared basic instructions, any such effect would work against the obtained pattern of results, in that it should reduce (not increase) the effect of stimulus relations in the control-instructions condition. Moreover, whether incentives were indeed necessary for the obtained pattern of results is an interesting question, but incentives are essential to rule out potential concerns that a null effect of control-instructions is due to insufficient motivation. Based on these considerations, the critical question is whether the obtained pattern of results is driven by (1) instructions not to be influenced by stimulus co-occurrence or (2) instructions to devote extra effort to processing stimulus relations (or both). Future research may help to address this question by orthogonally manipulating the two instruction components.

Conclusion

In sum, the current findings suggest that, although enhanced motivation to counteract effects of stimulus co-occurrence can strengthen the impact of stimulus relations on judgments and decisions, it does not reduce the impact of stimulus co-occurrence. This conclusion is consistent with dual-process theories of evaluative learning such as the APE model (Gawronski & Bodenhausen, 2011, 2018), but it is inconsistent with predictions derived from single-process propositional theories such as the IPM (De Houwer, 2018). Although more research is needed to investigate whether specific control strategies are more effective in reducing mere co-occurrence effects, the findings raise important questions about the limits of intentional control with significant implications for applied areas.

References

Aust, F., Diedenhofen, B., Ullrich, S., & Musch, J. (2013). Seriousness checks are useful to improve data validity in online research. *Behavior Research Methods, 45,* 527-535.

Bading, K., Stahl, C., & Rothermund, K. (2020). Why a standard IAT effect cannot provide evidence for association formation: The role of similarity construction. *Cognition and Emotion, 34,* 128-143.

Balas, R., & Gawronski, B. (2012). On the intentional control of conditioned evaluative responses. *Learning and Motivation, 43,* 89-98.

Chan, M. S., Jones, C. R., Jamieson, K. H., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of message countering misinformation. *Psychological Science, 28,* 1531-1546.

Corneille, O., & Hütter, M. (2020). Implicit? What do you mean? A comprehensive review of the delusive implicitness construct in attitude research. *Personality and Social Psychology Review, 24,* 212-232.

Corneille, O., & Mertens, G. (in press). Behavioral and physiological evidence challenges the automatic acquisition of evaluations. *Current Directions in Psychological Science*.

Corneille, O., Mierop, A., Stahl, C., & Hütter, M. (2019). Evidence suggestive of uncontrollable attitude acquisition replicates in an instructions-based evaluative conditioning paradigm: Implications for associative attitude acquisition. *Journal of Experimental Social Psychology, 85:103841*.

Corneille, O., & Stahl, C. (2019). Associative attitude learning: A closer look at evidence and how it relates to attitude model. *Personality and Social Psychology Review, 23,* 161-189.

Curran, P. J., & Hussong, A. M. (2009). Integrative data analysis: The simultaneous analysis of multiple data sets. *Psychological Methods, 14,* 81-100.

De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, *37*, 1-20.

De Houwer, J. (2011). Why the cognitive approach in psychology would profit from a functional approach and vice versa. *Perspectives on Psychological Science, 6,* 202-209.

De Houwer, J. (2018). Propositional models of evaluative conditioning. *Social Psychological Bulletin, 13*(3), e28046.

De Houwer, J., Van Dessel, P., & Moran, T. (2020). Attitudes beyond associations: On the role of propositional representations in stimulus evaluation. *Advances in Experimental Social Psychology, 61,* 127-183.

Gawronski, B., Balas, R., & Creighton, L. A. (2014). Can the formation of conditioned attitudes be intentionally controlled? *Personality and Social Psychology Bulletin, 40,* 419-432.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*, 692-731.

Gawronski, B., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model: Theory, evidence, and open questions. *Advances in Experimental Social Psychology, 44,* 59-127.

Gawronski, B., & Bodenhausen, G. V. (2015a). Theory evaluation. In B. Gawronski, & G. V. Bodenhausen (Eds.), *Theory and explanation in social psychology* (pp. 3-23). New York: Guilford Press.

Gawronski, B., & Bodenhausen, G. V. (2015b). Social-cognitive theories. In B. Gawronski, & G. V. Bodenhausen (Eds.), *Theory and explanation in social psychology* (pp. 65-83). New York: Guilford Press.

Gawronski, B., & Bodenhausen, G. V. (2018). Evaluative conditioning from the perspective of the associative-propositional evaluation model. *Social Psychological Bulletin, 13*(3), e28024.

Gawronski, B., & De Houwer, J. (2014). Implicit measures in social and personality psychology. In H. T. Reis, & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2nd edition, pp. 283-310). New York, NY: Cambridge University Press.

Gawronski, B., Mitchell, D. G. V., & Balas, R. (2015). Is evaluative conditioning really uncontrollable? A comparative test of three emotion-focused strategies to prevent the acquisition of conditioned preferences. *Emotion, 15,* 556-568.

Gawronski, B., Walther, E., & Blank, H. (2005). Cognitive consistency and the formation of interpersonal attitudes: Cognitive balance affects the encoding of social information. *Journal of Experimental Social Psychology*, *41*, 618-626.

Green, L. J. S., Luck, C. C., Gawronski, B., & Lipp, O. (in press). Contrast effects in backward evaluative conditioning: Exploring effects of affective relief/disappointment versus instructional information. *Emotion*.

Heycke, T., & Gawronski, B. (2020). Co-occurrence and relational information in evaluative learning: A multinomial modeling approach. *Journal of Experimental Psychology: General, 149,* 104-124.

Hu, X., Gawronski, B., & Balas, R. (2017). Propositional versus dual-process accounts of evaluative conditioning: I. The effects of co-occurrence and relational information on implicit and explicit evaluations. *Personality and Social Psychology Bulletin, 43,* 17-32.

Hütter, M., & Klauer, K. C. (2016). Applying processing trees in social psychology. *European Review of Social Psychology, 27,* 116-159.

Hütter, M., & Sweldens, S. (2018). Dissociating controllable and uncontrollable effects of affective stimuli on attitudes and consumption. *Journal of Consumer Research, 45,* 320-349.

Kukken, N., Hütter, M., & Holland, R. (2020). Are there two independent evaluative conditioning effects in relational paradigms? Dissociating effects of CS-US pairings and their meaning. *Cognition and Emotion, 34,* 170-187.

Kurdi, B., & Dunham, Y. (in press). Propositional accounts of implicit evaluation: Taking stock and looking ahead. *Social Cognition.*

Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest, 13,* 106-131.

Moors, A. (2016). Automaticity: Componential, causal, and mechanistic explanations. *Annual Review of Psychology, 67,* 263-287.

Moran, T., & Bar-Anan, Y. (2013). The effect of object–valence relations on automatic evaluation. *Cognition and Emotion*, *27*, 743-752.

Moshagen, M. (2010). multiTree: A computer program for the analysis of multinomial processing tree models. *Behavioral Research Methods, 42,* 42-54.

Nebenzahl, I. D., & Jaffe, E. D. (1998). Ethical dimensions of advertising executions. *Journal of Business Ethics, 17,* 805-817.

Noar, S. M., Francis, D. B., Bridges, C., Sontag, J. M., Ribisil, K. M., & Brewer, N. T. (2016). The impact of strengthening cigarette pack warnings: Systematic review of longitudinal observational studies. *Social Science and Medicine, 164,* 118-129.

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45*, 867-872.

Peters, K. R., & Gawronski, B. (2011). Are we puppets on a string? Comparing the impact of contingency and validity on implicit and explicit evaluations. *Personality and Social Psychology Bulletin, 37,* 557-569.

U.S. Food and Drug Administration (March 17, 2020). *FDA requires new health warnings for cigarette packages and advertisements*. Retrieved from https://www.fda.gov/news-events/press-announcements/fda-requires-new-health-warnings-cigarette-packages-and-advertisements?c (August 21, 2020).

Van Dessel, P., Gawronski, B., & De Houwer, J. (2019). Does explaining social behavior require multiple memory systems? *Trends in Cognitive Sciences, 23,* 368-369.

Zanon, R., De Houwer, J., & Gast, A. (2012). Context effects in evaluative conditioning of implicit evaluations. *Learning and Motivation*, *43*, 155-165.

Zanon, R., De Houwer, J., Gast, A., & Smith, C. T. (2014). When does relational information influence evaluative conditioning? *The Quarterly Journal of Experimental Psychology*, *67*, 2105-2122.

Table 1. Mean proportions and 95% confidence intervals of choice responses (yes vs. no) as a function of valence of co-occurring stimulus (positive vs. negative) and relation to co-occurring stimulus (stimulus causes vs. prevents co-occurring stimulus), integrative analysis of data from Experiments 1-3 (N = 1154).

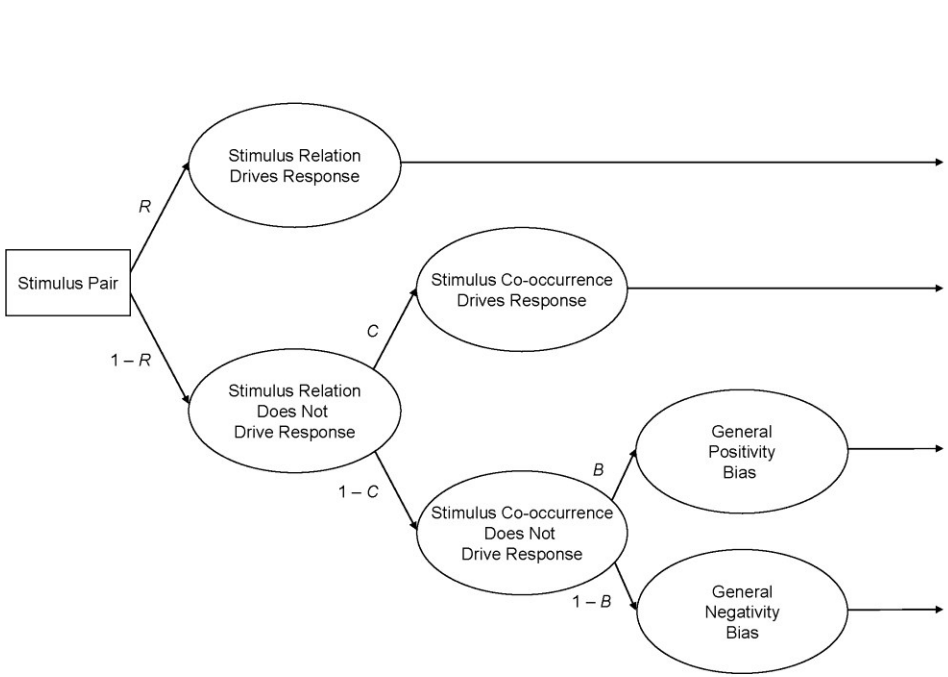
	Stimulus Causes Co-Occurring Stimulus		Stimulus Prevents Co-Occurring Stimulus	
	M	95% CI	*M*	95% CI
Standard Instructions				
Positive Co-Occurring Stimulus	.55	[.53, .57]	.42	[.40, .45]
Negative Co-Occurring Stimulus	.37	[.34, .39]	.47	[.45, .49]
Control Instructions				
Positive Co-Occurring Stimulus	.59	[.57, .61]	.45	[.42, .47]
Negative Co-Occurring Stimulus	.38	[.36, .40]	.53	[.51, .55]

Table 2. Parameter estimates without model restrictions as a function of instructions (standard-instructions vs. control-instructions), integrative analysis of data from Experiments 1-3 (N = 1154).

Parameter	Estimate	95% CI	$G^2(1)$	p	w
R					
Standard-instructions	.12	[0.10, 0.13]	271.58	< .001	.084
Control-instructions	.15	[0.14, 0.16]	434.05	< .001	.106
C					
Standard-instructions	.08	[0.06, 0.09]	98.60	< .001	.050
Control-instructions	.08	[0.06, 0.09]	81.08	< .001	.046
B					
Standard-instructions	.44	[0.43, 0.45]	192.56	< .001	.070
Control-instructions	.48	[0.47, 0.49]	25.91	< .001	.023

Note. The *R* parameter captures effects of stimulus relations; the *C* parameter captures effects of stimulus co-occurrence; the *B* parameter captures general response biases. G^2-values, p-values and effect sizes w refer to differences between parameter estimates and neutral reference points. The neutral reference point for *R* and *C* is 0; the neutral reference point for *B* is 0.5, with scores higher than 0.5 reflecting a general bias toward positive responses and scores lower than 0.5 reflecting a general bias toward negative responses.

Figure 1. Multinomial processing tree depicting effects of stimulus relations, stimulus co-occurrence, and general response biases on evaluative responses (positive vs. negative) for stimuli that cause or prevent either positive or negative stimuli. Adapted from Heycke and Gawronski (2020). Reprinted with permission.



Positive Co-occurring Stimulus		Negative Co-occurring Stimulus	
Stimulus Causes Co-occurring Stimulus	Stimulus Prevents Co-occurring Stimulus	Stimulus Causes Co-occurring Stimulus	Stimulus Prevents Co-occurring Stimulus
positive	negative	negative	positive
positive	positive	negative	negative
positive	positive	positive	positive
negative	negative	negative	negative