https://doi.org/10.1037/xge0001308

Does Co-Occurrence Information Influence Evaluations Beyond Relational Meaning? An Investigation Using Self-Reported and Mouse-Tracking Measures of Attitudinal Ambivalence

Jérémy Béna, Antoine Mauclet, & Olivier Corneille

UCLouvain, Belgium

Conflict of interest statement: The authors declare no conflict of interest.
Funding: This work was supported by an FNRS-FRS grant [grant T.0061.18] awarded to Olivier Corneille.

Author Note

Correspondence concerning this article should be addressed to Jérémy Béna, UCLouvain, PSP IPSY, 10 Place du Cardinal Mercier, 1348, Louvain-la-Neuve, Belgium. Email: jeremy.bena@uclouvain.be

Preregistrations, experiment programs, data, and analyses are available at https://osf.io/be9wt/

Parts of the results were presented at the *Journées de la Cognition Sociale*, Louvain-la-Neuve, Belgium, on October 14th, 2021, and at the *European Social Cognition Network (ESCON)* conference, Milano, Italy, on August 31st, 2022.

Acknowledgments

We thank Karoline Bading, Douglas Bancu, Olivier Desmedt, Doris Lacassagne, Martin Rouard, Christoph Stahl, and Colin Wathelet for their valuable comments during the preparation of the manuscript. We also thank Tobias Heycke and Christian Unkelbach for sharing their experimental materials.

CRediT authorship contribution statement

JB: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft. AM: Conceptualization, Formal analysis, Investigation, Methodology, Software, Writing – review & editing. OC: Conceptualization, Methodology, Supervision, Funding acquisition, Writing – original draft, Writing - review & editing.

Abstract

People occasionally encounter information whose structure bears divergent evaluative implications. For instance, when reading that a sunscreen protects against skin cancer, the relational meaning of the information (i.e., “protects against skin cancer”) has positive evaluative implications for the sunscreen, whereas the co-occurrence (of “sunscreen” with “skin cancer”) is negative. An important theoretical (and practical) issue is whether the co-occurrence information influences people’s evaluations beyond the relational meaning of the information. This question has been recently investigated using task comparison procedures (comparing evaluative outcomes on different tasks) and process dissociation procedures (estimating relational and co-occurrence parameters within a given task). In this article, we report four experiments that examined this question by reducing interpretational ambiguities inherent in the two preceding approaches. This was achieved by using self-reported and mouse-tracking measures of ambivalence. We reasoned that when co-occurrence and relational information have divergent (rather than convergent) evaluative implications, more ambivalence should be found. We tested this prediction in experiential (Experiments 1 to 3) and instructed (Experiment 4) procedures. Higher self-reported ambivalence was found in divergent than convergent conditions in all experiments. Ambivalence, as estimated with mouse-tracking measures, was higher in divergent than convergent conditions in the experiential experiments but not in the instructed experiment. Results are discussed with reference to single-process (propositional and episodic) and dual-process attitude learning models.

Keywords: attitudes; attitudinal ambivalence; dual-process theories of attitude; evaluative conditioning; mouse-tracking

Does Co-Occurrence Information Influence Evaluations Beyond Relational Meaning? An Investigation Using Self-Reported and Mouse-Tracking Measures of Attitudinal Ambivalence

People may come across information whose structure bears convergent or divergent evaluative implications. For instance, campaigns designed to reduce smoking may state that this habit causes cancer and display a picture of a tongue ulcer on a cigarette pack. In this case, the meaning of the information suggests that cigarettes should be negatively evaluated and the pairing of the cigarettes with a negative picture has *convergent* (here, negative) evaluative implications. Conversely, a campaign that states that vaccination prevents respiratory complications may display the picture of a patient under respiratory assistance. In this instance, the meaning of the information suggests that vaccines should be positively evaluated, but the pairing of the vaccines with a negative picture has a *divergent* (here, negative) evaluative implication. An important theoretical and practical question is whether, in cases like these, the co-occurrence information will influence the evaluation of the stimuli beyond the relational meaning of the information. This question is central to attitude models as it is concerned with whether there are learning processes that register co-occurrences between stimuli (i.e., an operating principle; see, e.g., Gawronski & Bodenhausen, 2009, 2014). Turning to more practical implications, if co-occurrence information influences evaluations beyond relational meaning, messages may be less effective when co-occurrence and relational meaning have divergent (e.g., "vaccination prevents respiratory complications") than convergent implications (e.g., "vaccination keeps lungs safe"). As a result, determining whether co-occurrences influence evaluations may prove critical in designing effective messages, for instance, in public health campaigns.

In the present research, we examined this question by relying on attitudinal ambivalence measures. The rationale is as follows: if co-occurrences influence evaluations over and above relational meaning, people should experience more mixed feelings (both positive and negative feelings) when these two types of information have divergent evaluative implications. In the remainder of the introduction, we explain how this question has been addressed in recent studies relying on task comparison and process dissociation procedures. We then discuss the value of ambivalence attitude measures to gain further insight into this question.

Investigating the influence of co-occurrence and relational meaning information in relational evaluative conditioning procedures

Recent research addressing the present question has relied on Evaluative Conditioning (EC) procedures. In EC procedures, a change in the evaluation of a Conditioned Stimulus (CS, a neutral stimulus) is typically observed after its pairing with a valent Unconditioned Stimulus (US; e.g., De Houwer 2007; Hofmann et al., 2010). For instance, pairing a CS (e.g., an unfamiliar brand) with a positive US (e.g., the picture of a puppy) typically results in a more positive evaluation of the CS.

In “Relational” EC procedures, CSs and USs are not merely paired; the relation between them is additionally specified. For instance, Förderer and Unkelbach (2012) showed participants portraits of males that were said to *love* or *loathe* animals. Moran et al. (2013) displayed alien creatures that *started* or *stopped* a pleasant melody or a human scream. Hu et al. (2017) presented pharmaceutical products that *caused* or *prevented* positive or negative health outcomes. When CSs love/start/cause USs, co-occurrence and relational meaning information have the same evaluative implications – as a result, they act in concert: they support convergent evaluative responses (e.g., positive if an alien creature starts a pleasant melody; negative if an alien creature starts a human scream). These are assimilative conditions. In contrast, when CSs loathe/stop/prevent USs, co-occurrence and relational meaning information have opposite evaluative implications. In this case, co-occurrence and relational information act in opposition: they support divergent evaluative responses. For instance, when a CS stops a positive US, the co-occurrence information supports a positive response, but the relational meaning information supports a negative response. These are contrastive conditions. Assimilative (e.g., “cause”) and contrastive (e.g., “prevent”) are highly reminiscent of inclusion and exclusion conditions of Process Dissociation Procedures (PDP; see Jacoby, 1991; Payne & Bishara, 2009; Yonelinas, 2012; see Hütter & Klauer, 2016 for a review in social psychology). We will return to this point in the “Process Dissociation” section.

Whether co-occurrence information influences CS evaluations beyond the relational meaning information has been tested using task comparison and process dissociation procedures. We discuss these procedures separately in the following two sections before going on to consider the value of an approach based on attitudinal ambivalence measures to examine this question.

Task comparison

In task comparison procedures (e.g., Hu et al., 2017; Moran et al., 2013, 2015, 2016), CS evaluations are compared on tasks assumed to reflect the contribution of more or less deliberate processes. Authors of task comparison procedure studies have found that CS evaluations are influenced by the relational information on self-reports and by the co-occurrence information on less deliberate evaluative tasks. For instance, Moran et al. (2013) found that participants preferred CSs that stopped negative USs to CSs that stopped positive USs on self-reports. Conversely, the same participants preferred CSs that stopped positive US to CSs that stopped negative US on an Implicit Association Test (Experiment 1) and a Sorting Paired Features task

(Experiment 2). Some studies demonstrated a main effect of co-occurrence information on direct evaluative tasks in addition to a main effect of relational information (e.g., Moran et al., 2016). These results, however, were not systematically observed; other task comparison studies failed to find an effect of co-occurrence above the effect of relational information (e.g., Hu et al., 2017, Experiment 3; Moran & Bar-Anan, 2019, Experiments 1 and 3; Zanon et al., 2014; see also Bading et al., 2020, for an alternative account of Moran et al.'s 2013 results).

Results from studies like these are interesting for a number of reasons. First, they suggest that co-occurrence information can influence evaluations. Even more remarkably, this was also the case when focusing the analyses on the subset of participants who correctly reported the relational meaning of the information in a subsequent memory task (Moran et al., 2013, 2016). This latter result makes it unlikely that the observed effects were obtained because the "loathe/stop/prevent" information was less correctly encoded or retrieved than the "love/start/cause" information. However, one limitation of these studies is that multiple processes underlying performance can differ between the two tasks. For instance, it may be the case that different stimulus dimensions were processed or recollected across tasks. For example, participants may have primarily retrieved the co-occurrence information when completing a speeded evaluative task such as an Implicit Association Test. Conversely, participants may have retrieved both co-occurrence and relational information when providing a non-speeded self-reported evaluation. In addition, task comparison outcomes do not directly address the joint influence of processes (sometimes acting in concert, sometimes in opposition) in attitude formation and evaluations. Yet, this joint activation is important for both theoretical and practical reasons. At a theoretical level, it may be interpreted as supporting evidence for the parallel operation of evaluative learning processes (see also the General Discussion). Theoretically and practically, only the joint operation of these processes may elicit attitudinal ambivalence.

Process dissociation

Process dissociation procedures address the same question using a rationale with an important advantage compared to the task comparison approach: no comparison between several tasks (e.g., between an indirect and a direct task) is involved. Instead, process dissociation studies estimate parameters (aimed at capturing psychological processes) by comparing performance in an inclusion condition and an exclusion condition *within a unique task*. The assumed processes act in concert in the inclusion condition (here, the assimilative condition) and in opposition in the exclusion condition. By applying a set of equations routinely implemented in Multinomial Processing Tree (MPT) models (see Calachini, 2020; Calanchini et al., 2018; Erdfelder et al., 2009; Riefer & Batchelder, 1988), estimates of the assumed parameters can be obtained. In relational EC procedures, the co-occurrence information and stimulus relation information are supposed to act in concert when CSs start/cause USs (inclusion condition) and they are supposed to act in opposition when CSs stop/prevent USs (contrastive condition).

Two equivalent MPT models were developed to disentangle the evaluative influences of co-occurrence and relational meaning information (Heycke & Gawronski, 2020; Kukken et al., 2020). In these models, stimulus relation determines evaluative responses with a probability *R* (for “relational” in Heycke & Gawronski’s RCB model; *m* for “meaning” in Kukken et al.’s model). In the absence of relational information, responses are based on co-occurrence information with a probability *C* (for “co-occurrence” in Heycke & Gawronski’s model; *p* for “pairing” in Kukken et al.’s model). In the absence of co-occurrence information, individuals guess “Positive” with a probability *B* (for “bias” in Heycke & Gawronski’s model; *g* for "guessing" in Kukken et al.'s model). MPT models of performance in Relational EC procedures were shown to fit the data well (Gawronski & Brannon, 2021; Heycke & Gawronski, 2020; Jin, Xia, Gawronski, & Hu, 2022; Kukken et al., 2020). Critically, both the R/m and C/p parameters were significantly above 0 in these studies, suggesting that both the relational meaning and the co-occurrence information contributed to evaluative responses (but see Gawronski, Luke, & Ng, 2021 for weak evidence for a C/p parameter above 0).

While process dissociation studies overcome interpretational ambiguities that characterize the task comparison approach, they also have limitations. First, a process dissociation approach does not allow examining the *joint* influence of the processes. This is because MPT models assume that *either* the relational information *or* the co-occurrence information will influence CS evaluations. Specifically, relational EC MPT models estimate responses based on stimulus relation (R/m) and responses based on co-occurrence information *in the absence of* stimulus relation (1-R/c*C/p). MPT models only capture the functional properties of the assumed processes. In other words, when stimulus relation drives a response (with a probability R/m), one does not need to model co-occurrence information as stimulus relation is sufficient to model responses. As a result, relational EC MPT models do not allow one to test whether both processes might sometimes be elicited at once (versus for different participants or for different evaluations). Second, and just as important, one advantage of the task comparison approach is lost because both correct and incorrect evaluative responses are necessarily involved in MPT modeling. Consequently, because there are interpretational ambiguities regarding incorrect trials (e.g., participants may have incorrectly encoded or retrieved the meaning of the stimuli), it is not clear whether the co-occurrence information parameter reflects an influence of CS-US pairings *despite* the correct interpretation of the CS-relation-US triplets. Validation studies (e.g., Kukken et al., 2020, which found different correlations between the R/m and C/m parameters and evaluative ratings) provide important empirical support for the notion that the parameters capture different processes. However, the important point here is that co-occurrence (as reflected in the C/p parameter) may come from trials for which participants incorrectly encoded the meaning of the triplets. In short, although incredibly valuable, process dissociation procedures are limited when it comes to gaining insights into whether co-occurrences influence evaluations. This is because relational EC MPT models (1) do not take into account the contribution of co-occurrence information when stimulus relation drives responses and (2) make use of trials for which participants may not have correctly encoded and retrieved the meaning of the triplets.

Using attitudinal ambivalence measures to gain insights into co-occurrence information effects

Attitudinal ambivalence is a state where an attitude object is associated with both positive and negative evaluations (Conner & Armitage, 2008; Schneider et al., 2021; Schneider & Schwarz, 2017; van Harreveld et al., 2015). Attitudinal ambivalence has not received much attention in the evaluative conditioning literature (for exceptions, see Glaser & Walther, 2012; Glaser et al., 2018; Rydell & McConnell, 2014). This may be considered surprising given that dual-process models of attitude posit the existence of multiple processes that can sometimes act in opposition. By definition, two processes acting in opposition support different responses. In contrastive conditions of relational EC procedures (e.g., a CS prevents a US), processes capturing co-occurrence information support a response in line with the US valence, while processes capturing stimulus relations support a response contrary to the US valence. If both classes of processes are elicited on a trial basis, attitudinal ambivalence should emerge as both positive and negative evaluations would be elicited.

Ambivalence should not be confused with neutrality (Rothman et al., 2017; Schneider et al., 2016, 2021; van Harreveld et al., 2015). Whereas neutrality is a state of both low positive and negative evaluations, ambivalence is a state of both high positive and negative evaluations. Ambivalence can be measured with several tools, including self-reported (e.g., scales of felt ambivalence; Priester & Petty, 1996) and mouse-tracking measures (e.g., Schneider et al., 2015, 2016), both of which are frequently used.

Ambivalence measures may prove helpful in addressing the (parallel) contribution of co-occurrence information beyond stimulus relation. If co-occurrence information and stimulus relation sometimes act in concert (i.e., in assimilative conditions, e.g., “cause” or “love”), and sometimes act in opposition (i.e., in contrastive conditions, e.g., “prevent” or “loathe”), ambivalence should be higher in contrastive than in assimilative conditions. This hypothesis cannot be tested with a task comparison approach or process dissociation studies. A critical asset of the new approach we advocate is that, when using ambivalence measures in a relational EC procedure, this hypothesis can be tested within a single evaluative task (overcoming limitations of task comparison procedures), and analyses can be restricted to correct trials only, for which there is a better guarantee of correct encoding and retrieval of the relational meaning of the information (thereby overcoming limitations of process dissociation procedures).

Overview of the experiments

We conducted four preregistered experiments. In Experiments 1 and 2, we relied on the procedure that Heycke and Gawronski have used (2020; see also Gawronski & Brannon, 2021; Hu et al., 2017; Jin et al., 2022). In the conditioning phase, participants saw hypothetical pharmaceutical products that caused or prevented positive or negative health conditions (resulting in four stimulus categories). Participants were then asked to evaluate the pharmaceutical products using their mouse (to record mouse trajectories) and to rate the ambivalence they felt about the pharmaceutical products. In Experiment 1, we used the same 12 CSs and USs Heycke and Gawronski have used. Trials were randomly displayed throughout the conditioning phase regardless of the specific relational information. The order of the evaluative tasks (mouse-tracking and felt ambivalence scales) was counterbalanced between participants.

Because Experiment 1 provided only weak evidence for attitudinal ambivalence in relational evaluative conditioning procedures, Experiment 2 was designed to increase the correct encoding/retrieval of both relational information and US Valence. To achieve this, we reduced the number of CSs and USs (from 12 to 8), and we blocked relational information in the conditioning phase.

We conducted Experiment 3 to address the limitations of Experiments 1 and 2 inherent in Heycke and Gawronski's (2020) conditioning phase (i.e., some relations may contradict existing mental models and explain at least part of the results of Experiments 1 and 2) and the method used to compute correct responses (i.e., in the mouse-tracking task, see Experiment 3 for more detail). In Experiment 3, we adapted the conditioning phase from Förderer and Unkelbach (2012): participants were presented with CSs (nonwords) said to "love" (assimilative condition) or "loathe" (contrastive condition) USs (negative or positive animals). We also added a role memory task at the end of the experiment and used responses to this task to compute the "correct" responses (see the presentation of Experiment 3 for more information).

In Experiments 1 to 3, participants experienced the CS-relation-US triplets during the conditioning phase. An important question is whether similar evidence for attitudinal ambivalence would still be found in an *instructed* procedure – that is, when participants are only instructed about the triplets but do not experience them directly. We will elaborate on this point when we introduce Experiment 4 below. In Experiment 4, we created an instruction-based version of Experiment 3 to investigate this question (see the presentation of Experiment 4 below).

In the four experiments, we additionally applied Heycke and Gawronski's RCB MPT model (2020; similar to Kukken et al.'s model, 2020) to the evaluative classifications collected in the mouse-tracking task. As mentioned above, the RCB model has allowed researchers to disentangle the contribution of co-occurrence and stimulus relation information in classification tasks by modeling response frequencies. By using (1) mouse-tracking measures, (2) felt ambivalence ratings, and (3) MPT parameter estimates, the present studies provide rich data on the contribution of co-occurrence information to evaluations.

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the four studies. Preregistrations, experiment programs, data, and analyses[1] of the four studies are publicly available on the Open Science Framework: https://osf.io/be9wt/. The

[1] We performed all analyses in R (R Core Team, 2020). We used *mousetrap* (Kieslich et al., 2019, version 3.2.0) to compute the mouse-tracking measures. We conducted the repeated-measures and mixed ANOVAs with *afex* (Singmann et al., 2017, version 1.0-1). We conducted pairwise multiple comparisons with *emmeans* (Lenth, 2020, version 1.5.2-1) and the effect sizes of these comparisons were calculated with *effectsize* (Ben-Schachar, Lüdecke, & Makowski, 2020). We fitted the RCB MPT model with *MPTinR* (Singmann & Kellen, 2013, version 1.14.1). We made the raincloud plots with *ggplot*2 (Wickham, 2016), *ggpubr* (Kassambara, 2020, version 0.4.0) and R scripts from Allen et al. (2021).

study project was approved by the ethics committee of the Psychological Sciences Research Institute at UCLouvain (Projet 2018-06).

Experiment 1

Participants and design

The design was a 2 (US Valence: Positive vs. Negative) × 2 (CS-US Relation: Cause vs. Prevent), with all factors within participants. At the evaluation stage, we counterbalanced the order of the tasks (evaluative classifications with mouse-tracking measures; felt ambivalence) between participants.

We recruited 240 participants (our targeted sample size) on Prolific. Participants (1) were English speakers, (2) were right-handed, (3) had an approval rate of at least 95%, and (4) had at least 100 previous submissions. Participants were paid US$3.01 for completing the study (for an estimated completion time of 17 minutes). We excluded 17 participants (two did not end the study, resulting in no data; seven declared they did not pay attention to the stimuli or that they did not take their responses seriously; five had mean response times below 300 milliseconds or above 5000 milliseconds in the mouse-tracking task; we were not able to retrieve the mouse-tracking data of three participants). This resulted in a final sample size of 223 participants (67.26% female, three not reported; M_{age} = 37.52; SD_{age} = 12.35, two not reported).

To determine sample size, we set α to .05, and we aimed for a statistical power of 80% to detect an effect as small as Cohen's *d* = 0.2 in a paired samples *t*-test. An analysis with G*Power (Version 3.1.9.6, Faul et al., 2007) showed that we would need 199 participants. To avoid a final sample smaller than the targeted sample size, we increased this estimate by 20%, resulting in a targeted sample size of 240 participants.

Materials and procedure

We programmed the experiment with the beta version of lab.js (Henninger et al., 2021) and the Mousetrap plugin (Henninger et al., 2020). We used JATOS (Lange, Kühn, & Filevich, 2015) to run the study online on Prolific.

Stimuli

We used the 12 CSs, 12 USs (half positive), and the conditioning phase of Heycke and Gawronski (2020, Experiment 1). The 12 CSs are images of hypothetical pharmaceutical products (e.g., “Shimeron”). The USs depict health conditions (USs+: positive health conditions, e.g., a happy old couple; USs-: negative health conditions, e.g., a skin rash).

Conditioning phase

Participants were told that the study investigated how people form impressions of novel objects. During the conditioning phase, each US was repeatedly paired with one CS. On each trial, the CS-US pair was displayed together with relational information qualifying their link: CSs were said to either cause or prevent the USs. CSs always appeared on the left side, USs appeared on the right side, and relational information appeared in the center of the screen. The stimuli were presented simultaneously on the screen for 3 seconds, with a 1-second inter-trial interval with a fixation cross. For each participant, each CS was allocated to one US Valence and one CS-US Relation condition by means of a Latin square. The conditioning phase was divided into four blocks of 24 trials each, resulting in a total of 96 trials (each CS-relation-US trial was displayed eight times in total, two times in each block). After reading about the four types of trials that participants would see, participants were instructed to think of the image pairs in terms of the relation mentioned on the screen (“causes” or “prevents”). Between blocks, participants knew their progression in the conditioning phase (e.g., after the first block, participants read:

"This is the end of the first block. You have finished 25% of the task. Press the spacebar to continue with the next block") and could take self-paced pauses.

Evaluation stage

After the conditioning phase, participants entered the evaluation stage, consisting of two tasks (whose order was counterbalanced between participants): a felt ambivalence task and an evaluative classification task with mouse-tracking measures (called the mouse-tracking task hereafter).

Felt ambivalence task.

In the felt ambivalence task (Priester & Petty, 1996), the 12 CSs displayed in the conditioning phase were presented once individually in a random order. Participants used 11-point Likert scales to indicate how much they felt (1) conflicted (0: "Feel no conflict at all"; 10: "Feel maximum conflict"), (2) indecision (0: "Feel no indecision at all"; 10: "Feel maximum indecision"), and (3) mixed about the products (0: "I am completely one-sided"; 10: "I have completely mixed reactions"). The instruction read: "*In the following part, we are interested in how you feel about the products of the preceding task. Please indicate your personal feeling on each product by clicking on the response option that best reflects your personal evaluation on each question*." As the internal reliability of the three items was high (Cronbach's α = .88), the felt ambivalence scores were the participants' average response across the three items (the higher the scores, the more ambivalent participants felt).

Evaluative classification task with mouse-tracking measures.

In the mouse-tracking task, each trial began with participants clicking on a "Start" button at the bottom of the screen. Upon clicking, one of the 12 CSs appeared on the screen (in a random order). Participants were instructed to use their mouse to click on one of two responses displayed at each corner at the top of the screen: “Positive” or “Negative” (response position was counterbalanced between participants). Each CS was presented once. The instruction read as follows: *“In the following part, we are interested in how you would evaluate the products of the preceding task. Please indicate your personal evaluation of each product by clicking on the response option that best reflects your personal evaluation. A typical trial will proceed as follows: The evaluative response options will appear at the top of the computer screen. As soon as you click on the “Start” button, the picture of one of the products will appear on the screen. It is important to begin moving the mouse toward a response button immediately after the picture is displayed. As soon as you have clicked on a response, the next trial will begin. Please remember to begin moving the mouse as soon as you see the picture of a product.”*

Check measures and socio-demographic information.

After the evaluation stage, participants indicated whether they used a mouse (n = 97), a touchpad/trackpad (n = 126), or another pointing device (n = 0); their age and gender; and any comment they may have. Then, they answered an attention check similar to the one used by Heycke and Gawronski (2020): *Did you pay attention to the images presented throughout the entire task? (the response to this question will not affect your payment)* (answer: “Yes” or “No”). Participants also answered the following seriousness check (based on Aust et al., 2013): *“It would be very helpful if you could tell us at this point whether you have taken the requested responses seriously, so that we can use your answers for our scientific analysis, or whether you were just clicking through to take a look at the survey? (this will not affect your payment).”* (answer: “I have taken the requested responses seriously” or “I have just clicked through, please discard my data”). We used the answers to the attention and seriousness checks as exclusion criteria (see the Design and participants section). Participants were then thanked and debriefed.

Results

Preregistered analyses

We first report the preregistered analyses on all trials from the felt ambivalence (in which we calculated mean felt ambivalence scores) and mouse-tracking tasks (in which we calculated Maximum Deviation scores, MD, and Area Under the Curve scores, AUC). Importantly, we also report preregistered analyses only on the subset of CSs that participants correctly evaluated in the mouse-tracking task. In correct trials, participants reported the correct meaning of the CS-US pairs: When CSs cause USs- and USs+, the correct meaning of the pairings is negative and positive, respectively. When CSs prevent USs- and USs+, the correct meaning of the pairings is positive and negative, respectively. For this subset of correct trials, we reasoned that there is evidence that participants correctly interpreted and retrieved the meaning of the pairs. As a result, finding more ambivalence in the Prevent vs. Cause condition would suggest that processes act in opposition in the Prevent condition but in concert in the Cause condition. Analyses on the correct trials only are based on a smaller sample size than analyses on all trials, as each participant had to have two (in the repeated-measures *t*-tests) and four (in the repeated-measures ANOVAs) observations. Because some participants only provided incorrect responses in at least one of the cells, they were excluded from analyses on the correct trials.

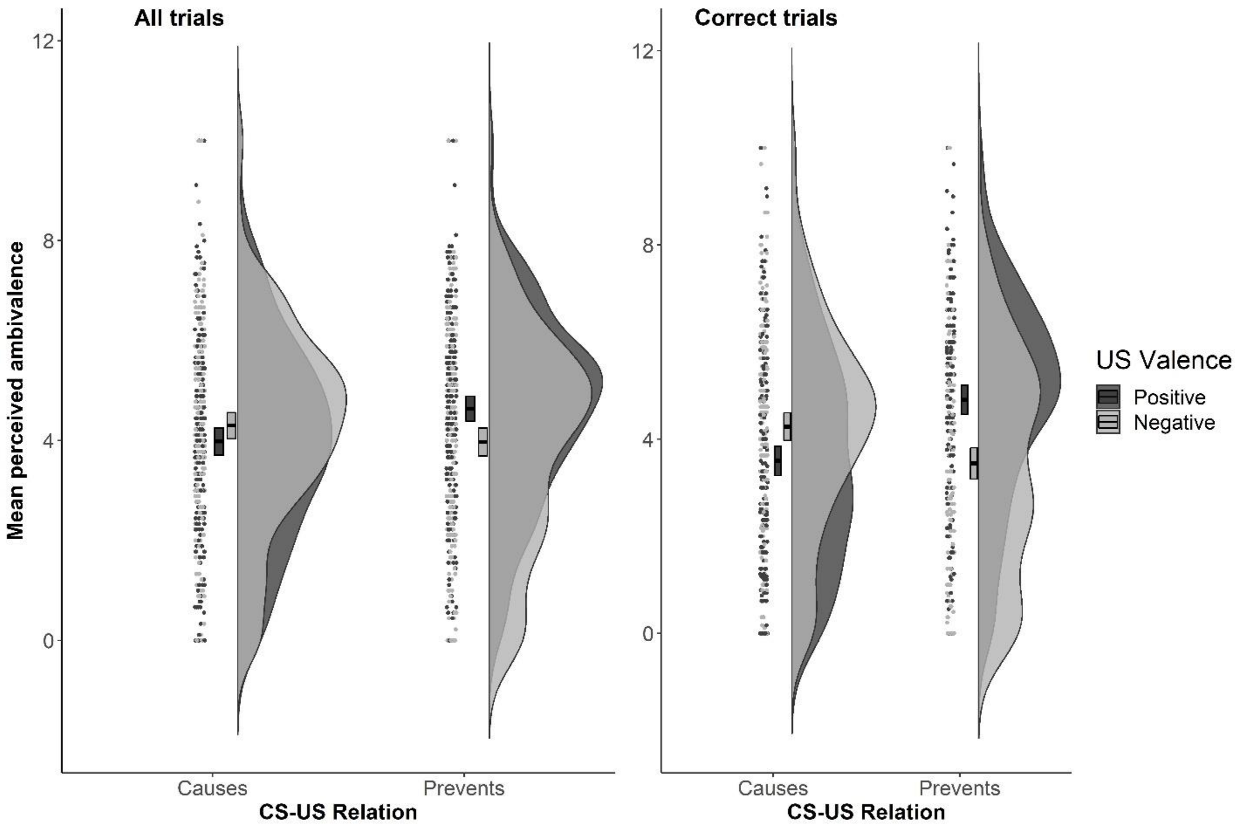

Figure 1. Mean felt ambivalence as a function of CS-US Relation and US Valence on all trials (left) and only on trials where participants provided the correct meaning-based response (right) in Experiment 1. The dots are the participants' scores (jittered). The lower and upper limits of the boxplots are the 95% confidence intervals, with the mean in between. The distributions represent the kernel probability density of the data.

All trials

Felt ambivalence scores.

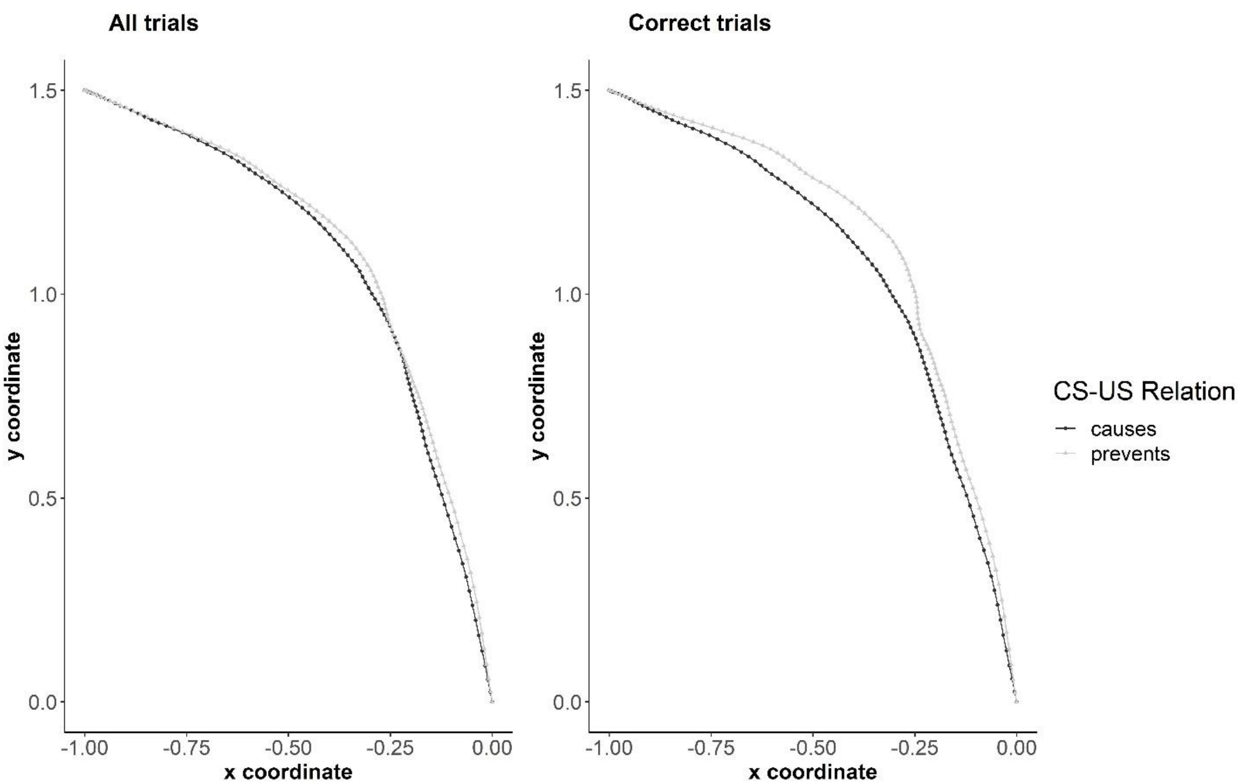

Figure 2. Average mouse trajectories as a function of CS-US Relation on each trials (left panel) and only on trials where participants provided the meaning-based correct response (right panel) in Experiment 1. All trajectories are mapped to the left.

We conducted a 2 (US Valence: Positive vs. Negative) × 2 (CS-US Relation: Cause vs. Prevent) repeated-measures ANOVA on felt ambivalence scores. The results are displayed in Figure 1 (left panel) and Table 1. We found no main effect of US Valence or of CS-US Relation, $F(1, 222) = 2.45$, $p = .119$, $\eta^2_G = .002$; $F(1, 222) = 3.71$, $p = .055$, $\eta^2_G = .002$, respectively. The two-way interaction was significant, $F(1, 222) = 28.6$, $p < .001$, $\eta^2_G = .015$. For CSs paired with positive USs, felt ambivalence scores were higher in the Prevent than in the Cause condition, $t(222) = -5.35$, $p < .0001$, $d = 0.36$, while the opposite was true for CSs paired with negative USs: felt ambivalence scores were higher in the Cause than in the Prevent condition, $t(222) = 2.56$, $p = .011$, $d = 0.17$.

Table 1. *Mean (and SD) scores in the felt ambivalence and mouse-tracking tasks as a function of US Valence and CS-US Relation in Experiment 1*

	Felt ambivalence		Mouse-tracking			
			AUC		MD	
	All	Correct	All	Correct	All	Correct
US Valence: Positive						
Cause	3.98 (2.06)	3.39 (2.19)	.35 (.28)	.33 (.34)	.44 (.37)	.42 (.43)
Prevent	4.64 (1.9)	4.7 (2.16)	.37 (.31)	.45 (.44)	.5 (.4)	.58 (.57)
US Valence: Negative						
Cause	4.3 (1.98)	4.06 (2.1)	.4 (.33)	.41 (.37)	.52 (.41)	.57 (.49)
Prevent	3.97 (2.13)	3.38 (2.29)	.39 (.32)	.43 (.45)	.5 (.41)	.57 (.54)

Note. AUC: Area Under the Curve; MD: Maximum Deviation.

Mouse-tracking measures.

In the mouse-tracking task, we computed two measures of curvature: the Maximum Deviation (MD; the maximum absolute deviation of the trajectory from the idealized direct path) and the Area Under the Curve (AUC; the geometric area between the trajectory and the idealized direct path). The two measures were highly correlated, $r(221) = .906$, $p < .001$, but we performed separate analyses on the two scores. Figure 2 (left panel) displays the mean mouse trajectories as a function of CS-US Relation.

We conducted paired samples *t*-tests on MD and AUC scores with CS-US Relation as a factor. The effect of CS-US Relation was not significant in both tests, $t(222) = -0.84$, $p = .403$, $d = 0.06$ (MD scores); $t(222) = -0.448$, $p = .655$, $d = 0.03$ (AUC scores).

Complementarily, we conducted a 2 (US Valence: Positive vs. Negative) × 2 (CS-US Relation: Cause vs. Prevent) repeated-measures ANOVAs on MD and AUC scores (see Table 1 for the descriptive statistics). We found no significant main or interactive effect on these mouse-tracking measures (MD: *F*s(1, 222) ≤ 3.33, *p*s ≥ .07, $\eta^2_G \leq .003$; AUC: *F*s(1, 222) ≤ 2.47, *p*s ≥ .118, $\eta^2_G \leq .002$).

RCB MPT model

We computed the aggregated frequencies of “Positive” and “Negative” responses at each CS-US Relation × US Valence level. This allowed us to apply the RCB Multinomial Processing Tree (MPT) model of Heycke and Gawronski (2020; see Kukken et al., 2020 for an analogous model). The RCB model contains three parameters to account for performance in the mouse-tracking task (where participants were asked to provide evaluative classifications of each CS). Participants can evaluate the CSs based on the CS-US Relation (captured by the *R* parameter). In the absence of relational information (1-*R*), participants may respond based on the co-occurrence between CSs and USs (captured by the *C* parameter). Finally, in the absence of co-occurrence information (1-*C*), participants may guess that the CS is positive (captured by the *B* parameter). For the main results, see Table 2 (see Table 3 for the proportions of evaluative classifications). Overall, the model did not fit the data well, $G^2(1) = 4.72$, $p = .03$ (see the additional, non-preregistered analyses below for model fit and parameter estimates in each task order condition at the evaluation stage).

Table 2. *Parameter estimates and goodness-of-fit statistics of Heycke and Gawronski's RCB MPT model in Experiments 1 to 4*

	Parameter estimates (95% CI)			Model fit	
	R	C	B	$G^2(1)$	*p*-value
Experiment 1 - All	.31 (.27, .35)	.17 (.11, .22)	.5 (.47, .53)	4.72	0.03
- *Mouse-tracking first*	.33 (.28, .38)	.15 (.07, .23)	.51 (.46, .55)	0.27	0.6
- *Mouse-tracking second*	.29 (.23, .34)	.18 (.11, .25)	.5 (.45, .54)	6.45	0.01
Experiment 2	.4 (.36, .44)	.38 (.31, .45)	.5 (.45, .56)	1.55	0.21
Experiment 3	.15 (.1, .19)	.24 (.19, .3)	.6 (.57, .64)	102.55	< .001
Experiment 4	.17 (.12, .21)	.18 (.13, .24)	.52 (.49, .56)	63.87	< .001

Note. R: Probability that the evaluative response is based on the relational information; C: Probability that the evaluative response is based on co-occurrence information; B: Probability to guess “Positive”. A *p*-value below the alpha threshold is indicative of a bad model fit.

Correct trials only

Felt ambivalence scores.

We repeated the 2 (US Valence: Positive vs. Negative) × 2 (CS-US Relation: Cause vs. Prevent) repeated-measures ANOVA on felt ambivalence scores now on the more diagnostic correct trials only (see Figure 1, right panel; Table 1; for the mean proportions of “positive” responses, see Table 3). Contrary to the results found when collapsing across correct and incorrect trials, we found a significant main effect of CS-US Relation, $F(1, 161) = 5.11$, $p = .025$, $\eta^2_G = .005$, even though the result is just below the alpha threshold and the effect size is very small. Felt ambivalence scores were higher in the Prevent

Table 3. *Mean (and SD) proportions of "positive" responses in the mouse-tracking task as a function of US Valence and CS-US Relation in Experiments 1 and 2*

	Experiment 1	Experiment 2
US Valence: Positive		
Cause	0.7 (0.29)	0.82 (0.28)
Prevent	0.42 (0.34)	0.4 (0.4)
US Valence: Negative		
Cause	0.27 (0.28)	0.2 (0.31)
Prevent	0.62 (0.33)	0.57 (0.39)

condition than in the Cause condition. The main effect of US Valence was significant too, $F(1, 161) = 4.56$, $p = .034$, $\eta^2_G = .006$. The interaction was significant again, $F(1, 161) = 47.32$, $p < .001$, $\eta^2_G = .049$. For CSs paired with positive USs, felt ambivalence scores were higher in the Prevent than in the Cause condition, $t(161) = -6.53$, $p < .0001$, $d = 0.51$. The opposite was true for CSs paired with negative USs; here, felt ambivalence scores were higher in the Cause than in the Prevent condition, $t(161) = 3.43$, $p = .0008$, $d = 0.27$.

Mouse-tracking measures.

Figure 2 (right panel) displays the mean mouse trajectories as a function of CS-US Relation. We conducted paired samples *t*-tests on MD and AUC scores with the CS-US Relation as a factor. The predicted effect of CS-US Relation was significant on MD scores, $t(217) = -2.00$, $p = .047$, $d = 0.14$, but not on AUC scores, $t(217) = -1.73$, $p = .085$, $d = 0.12$. Again, the results are very close to the alpha threshold, and the effect sizes are very small.

Complementarily, we conducted a 2 (US Valence: Positive vs. Negative) × 2 (CS-US Relation: Cause vs. Prevent) repeated-measures ANOVAs on MD and AUC scores (see Table 1 for the descriptive statistics). We found the predicted significant main effect of CS-US Relation on both MD and AUC scores, $F(1, 161) = 4.79$, $p = .03$, $\eta^2_G = .006$ and $F(1, 161) = 6.13$, $p = .014$, $\eta^2_G = .007$, respectively. The main effect of US Valence was not significant, $F(1, 161) = 3.38$, $p = .068$, $\eta^2_G = .004$ (MD scores) and $F(1, 161) = 1.26$, $p = .263$, $\eta^2_G = .001$ (AUC scores).

The two-way interaction was significant on MD scores, $F(1, 161) = 4.48$, $p = .036$, $\eta^2_G = .007$, but not on AUC scores, $F(1, 161) = 1.92$, $p = .167$, $\eta^2_G = .004$. For CSs paired with positive USs, MD scores were higher in the Prevent than in the Cause condition, $t(161) = -3.09$, $p = .002$, $d = 0.24$, while there was no significant effect of CS-US Relation for CSs paired with negative USs, $t(161) = 0.11$, $p = .917$, $d = .008$.

Additional, non-preregistered analyses

RCB MPT model

In the preregistered analyses, we found that the RCB MPT model did not fit the data well. In this analysis, however, we aggregated response frequencies from participants who performed the mouse-tracking task first or second (i.e., after the felt ambivalence task). Performing the mouse-tracking task second might have decreased the overall model fit because of carry-over effects of felt ambivalence ratings. To explore whether this was the case, we tested the model fit in each task's order of the evaluation stage (mouse-tracking first or second). Table 2 summarizes the main results. The model did not fit the data well when participants performed the mouse-tracking task second, $G^2(1) = 6.45$, $p = .011$. When participants performed the mouse-tracking task first, however, the model fitted the data well, $G^2(1) = 0.27$, $p = .6$. The *R* parameter (capturing the probability to give a response based on the relational information) was significantly above 0, as indicated by a decrease in the model fit when *R* is restricted to 0, $\Delta G^2(1) = 155.31$, $p < .001$. The *C* parameter (capturing the probability to give a response based on co-occurrence information) was also significantly above 0, $\Delta G^2(1) = 15.07$, $p < .001$. The *B* parameter (guessing "positive") was not different from .5, $\Delta G^2(1) = 0.06$, $p = .8$.

Discussion

Experiment 1 provides initial support for the contribution of both relational and co-occurrence information on felt and mouse-tracking measures of ambivalence. Of note, effects were best observed when focusing the analyses on correct trials only, which greatly reduced interpretational ambiguities (as these trials are likely associated with a correct understanding and recollection of the relational information). On the mouse-tracking measure, we found a larger maximal departure to the ideal trajectory for "prevent" than for "cause" CSs. In the correct trials, this indicates that participants shifted their mouse toward co-occurrence information. Likewise, participants felt more mixed about these CSs, suggesting that opposing processes were at play.

Nevertheless, it is worthwhile noting that the effects on the self-reported measures were qualified by the US valence (we will return to this point in the General Discussion). Furthermore, effects were generally weak and close to the .05 threshold. Finally, the MPT model fitted the data well only when the self-reported measures were collected after the evaluative classification measures.

Experiment 2

Experiment 2 was intended to test the robustness of the effects under conditions that were possibly more conducive to co-occurrence information effects. We reasoned that presenting the relational information in blocks in the conditioning phase may better emphasize the co-occurrence information. In two consecutive blocks, the CSs either always caused or prevented the USs, so only the US valence varied within the blocks. We surmised that this adapted procedure might increase attention to the US valence information, which, in turn, could facilitate its evaluative influence over and above that of the relational meaning of the pairing. To facilitate the encoding/retrieval of the relational information, we also reduced the number of CS-US pairs (from 12 in Experiment 1 to eight here).

Participants and design

The design was a 2 (US Valence: Positive vs. Negative) × 2 (CS-US Relation: Cause vs. Prevent), with all factors within participants. In the conditioning phase, we counterbalanced the presentation order of the relational information (cause first; prevent first) between participants.

Following the same rationale as Experiment 1, we recruited 240 participants (our targeted sample size) on Prolific. Again, participants (1) were English speakers, (2) were right-handed, (3) had an approval rate of at least 95%, and (4) had at least 100 previous submissions. Participants did not take part in Experiment 1. Participants were paid US$2.48 for completing the study (for an estimated completion time of 14 minutes). We excluded 23 participants (three did not end the study, resulting in no data; six declared they did not pay attention to the stimuli or that they did not take their responses seriously; 14 had mean response times below 300 milliseconds or above 5000 milliseconds in the mouse-tracking task; we were not able to retrieve the mouse-tracking data of one participant – participants might have been excluded regarding more than one criterion). This resulted in a final sample size of 217 participants (63.59% female, four not reported; M_{age} = 35.38; SD_{age} = 11.83, three not reported).

Materials and procedure

Experiment 2 departed from Experiment 1 in three regards. First, we reduced the number of stimuli and trials by randomly selecting eight CSs and eight USs (four positive; four negative) in the materials of Experiment 1 – resulting in a total of 64 trials. Second, the conditioning phase was divided into two blocks of 32 trials each. In each block, only one CS-US Relation condition

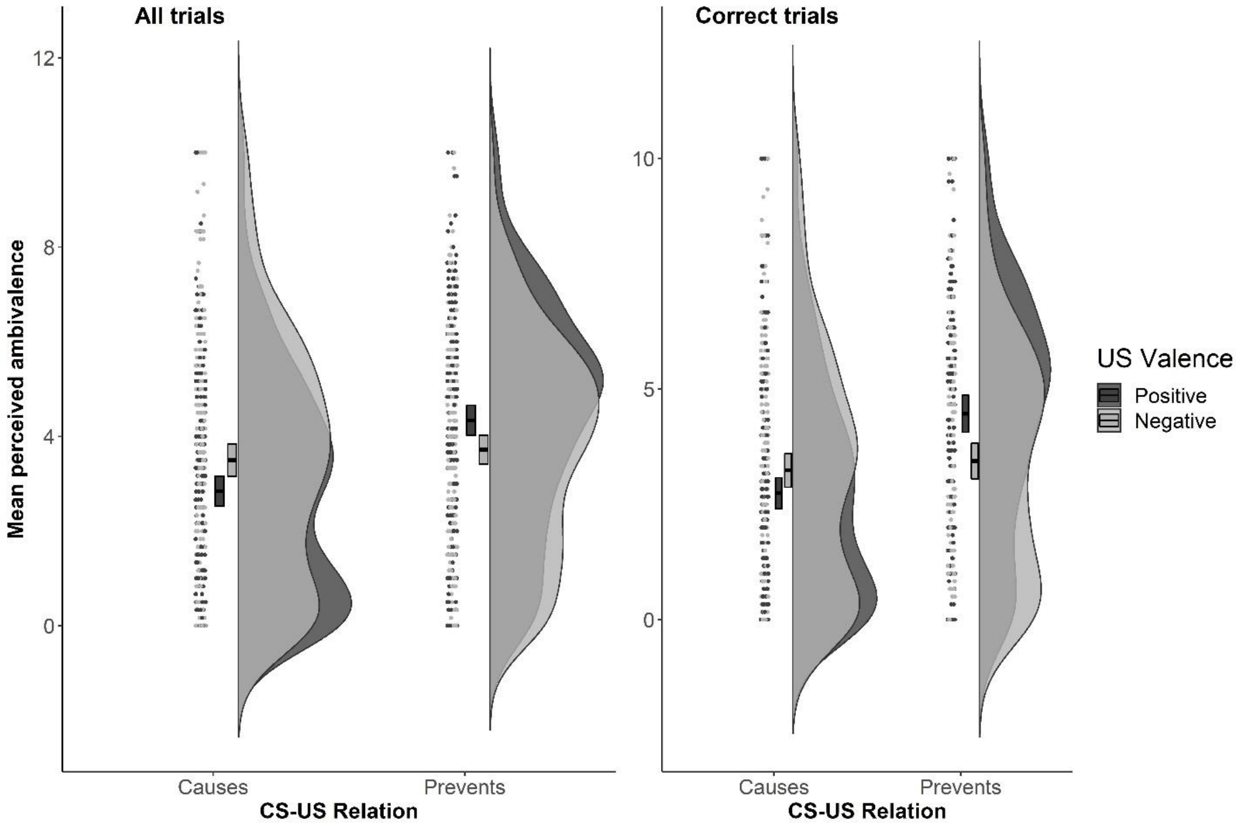

Figure 3. Mean felt ambivalence as a function of CS-US Relation and US Valence on all trials (left) and only on trials where participants provided the correct meaning-based response (right) in Experiment 2. The dots are the participants' scores (jittered). The lower and upper limits of the boxplots are the 95% confidence intervals, with the mean in between. The distributions represent the kernel probability density of the data.

(either Cause or Prevent) was displayed. Within blocks, the US Valence varied. The blocks' order was counterbalanced between participants. Between the two blocks, participants could take a self-paced pause. Third, at the evaluation stage, we decided to have participants complete the mouse-tracking measures always first. This allowed avoiding carry-over effects on both mouse-tracking ambivalence measures and MPT modeling derived from the mouse-tracking task. In addition, effects on the felt ambivalence measures were not qualified by an order effect in Experiment 1.

Results

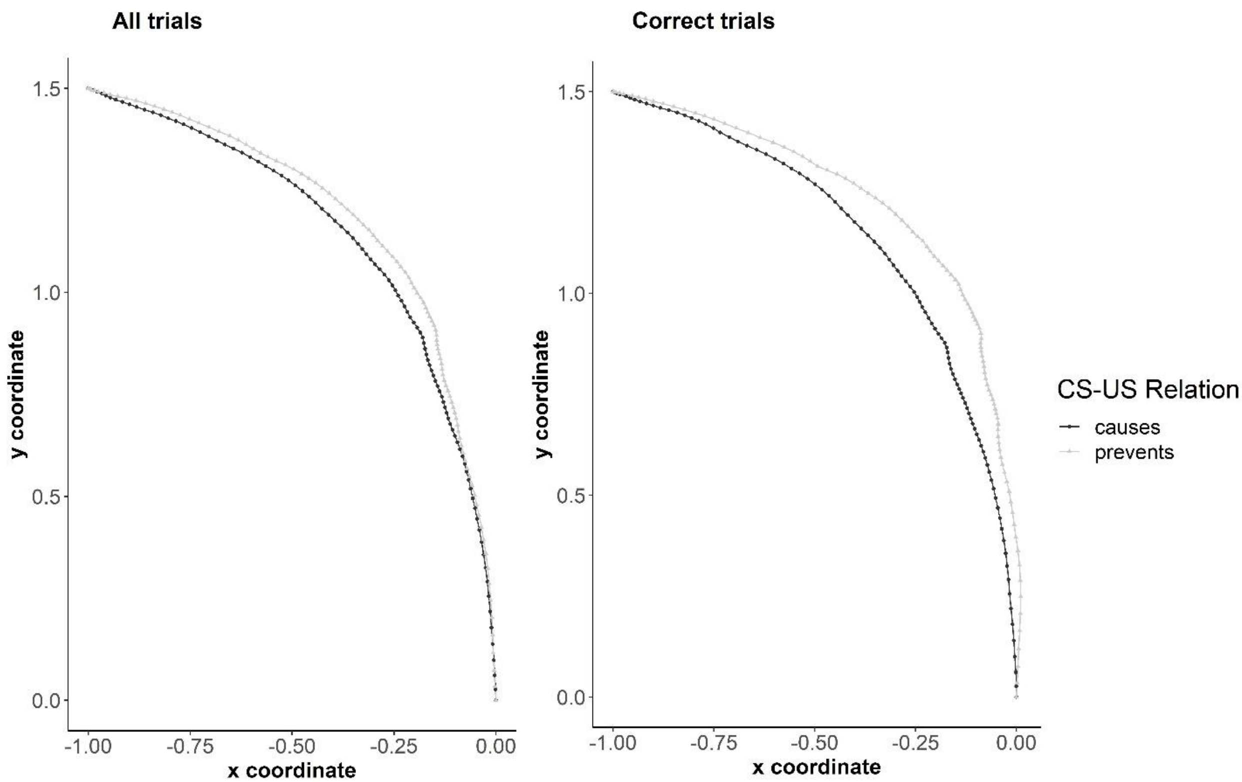

Figure 4. Average mouse trajectories as a function of CS-US Relation on each trial (left panel) and only on trials where participants provided the meaning-based correct response (right panel) in Experiment 2. All trajectories are mapped to the left.

Preregistered analyses

All trials

Felt ambivalence scores.

We conducted a 2 (US Valence: Positive vs. Negative) × 2 (CS-US Relation: Cause vs. Prevent) × 2 (Blocks' Order: Cause first vs. Prevent first) mixed ANOVA on felt ambivalence scores. The results are displayed in Figure 3 (left panel) and Table 4. We found no main effect of US Valence, $F(1, 215) = 0.02$, $p = .879$, $\eta^2_G < .002$. Contrary to Experiment 1, we found a significant main effect of CS-US Relation on all trials, $F(1, 215) = 38.39$, $p < .001$, $\eta^2_G = .031$: Felt ambivalence scores were higher in the Prevent condition than in the Cause condition. The

Table 4. *Mean (and SD) scores in the felt ambivalence and mouse-tracking tasks as a function of US Valence and CS-US Relation in Experiment 2*

	Felt ambivalence		Mouse-tracking			
			AUC		MD	
	All	Correct	All	Correct	All	Correct
US Valence: Positive						
Cause	2.85 (2.35)	2.44 (2.41)	.42 (.41)	.46 (.41)	.55 (.44)	.56 (.49)
Prevent	4.34 (2.39)	4.37 (2.58)	.44 (.34)	.47 (.38)	.58 (.44)	.65 (.52)
US Valence: Negative						
Cause	3.5 (2.56)	2.99 (2.71)	.41 (.33)	.47 (.36)	.55 (.44)	.6 (.49)
Prevent	3.72 (2.28)	3.22 (2.58)	.46 (.36)	.56 (.48)	.67 (.52)	.8 (.6)

Note. AUC: Area Under the Curve; MD: Maximum Deviation.

two-way interaction between US Valence and CS-US Relation was significant too, $F(1, 215) = 25.23$, $p < .001$, $\eta^2_G = .017$, but the pattern of simple effects was different from the one observed in Experiment 1. For CSs paired with positive USs, felt ambivalence scores were higher in the Prevent than in the Cause condition, $t(215) = -7.39$, $p < .0001$, $d = 0.5$, while no difference was found for CSs paired with negative USs, $t(215) = -1.31$, $p = .191$, $d = 0.09$. No effect including Blocks' Order was significant, $Fs(1, 215) \leq 2.84$, $ps \geq .093$, $\eta^2_G \leq .002$.

Mouse-tracking measures.

Figure 4 (left panel) displays the mean mouse trajectories as a function of CS-US Relation. We conducted paired samples *t*-tests on MD and AUC scores with the CS-US Relation as a factor. MD scores were higher in the Prevent than in the Cause condition, $t(216) = -2.58$, $p = .011$, $d = 0.18$. However, the effect of CS-US Relation was not significant on AUC scores, $t(216) = -1.46$, $p = .146$, $d = 0.1$.

Complementarily, we conducted a 2 (US Valence: Positive vs. Negative) × 2 (CS-US Relation: Cause vs. Prevent) × 2 (Blocks' Order: Cause first vs. Prevent first) mixed ANOVA on MD and AUC scores (see Table 4 for the descriptive statistics). Only the above-mentioned main effect of CS-US Relation on MD scores was significant, $F(1, 215) = 6.69$, $p = .01$, $\eta^2_G = .007$. All other effects were not significant, $F(1, 215) \leq 3.22$, $ps \geq .074$, $\eta^2_G \leq .003$.

RCB MPT model

The model fitted the data well, $G^2(1) = 1.55$, $p = .21$ (see Table 2 for the parameter estimates; see Table 3 for the proportions of evaluative classifications). The *R* parameter (capturing the probability to give a response based on the relational information) was significantly above 0, as indicated by a decrease in the model fit when *R* is restricted to 0, $\Delta G^2(1) = 307.19$, $p < .001$. The *C* parameter (capturing the probability to give a response based on co-occurrence information) was also significantly above 0, $\Delta G^2(1) = 108.65$, $p < .001$. The *B* parameter (guessing "positive") was not different from .5, $\Delta G^2(1) < 0.01$, $p = .99$.

Correct trials only

Felt ambivalence scores.

We repeated the 2 (US Valence: Positive vs. Negative) × 2 (CS-US Relation: Cause vs. Prevent) × 2 (Blocks' Order: Cause first vs. Prevent first) mixed ANOVA on felt ambivalence scores now on the correct trials only (see Figure 3, right panel; for the mean proportions of "positive" responses, see Table 3; see Table 4). Again, we found a significant main effect CS-US Relation, $F(1, 120) = 29.94$, $p < .001$, $\eta^2_G = .043$. The main effect of US Valence was not significant, $F(1, 120) = 2.78$, $p = .098$, $\eta^2_G = .003$. Again, the interaction between US Valence and CS-US Relation was significant, $F(1, 120) = 12.9$, $p < .001$, $\eta^2_G = .023$. For CSs paired with positive USs, felt ambivalence scores were higher in the Prevent than in the Cause condition, $t(120) = -6.22$, $p < .0001$, $d = 0.57$, while no difference was found for CSs paired with negative USs, $t(120) = -1.01$, $p = .314$, $d = 0.09$. No effect including Blocks' Order was significant, $Fs(1, 120) \leq 2.2$, $ps \geq .141$, $\eta^2_G \leq .004$.

Mouse-tracking measures.

Figure 4 (right panel) displays the mean mouse trajectories as a function of CS-US Relation. We conducted paired samples *t*-tests on MD and AUC scores with the CS-US Relation as a factor. The effect of CS-US Relation was now significant on both MD scores, $t(192) = -3.54$, $p = .0005$, $d = 0.25$ and AUC scores, $t(192) = -2.21$, $p = .028$, $d = 0.16$.

Complementarily, we conducted a 2 (US Valence: Positive vs. Negative) × 2 (CS-US Relation: Cause vs. Prevent) × 2 (Blocks' Order: Cause first vs. Prevent first) mixed ANOVA on MD and AUC scores (see Table 4 for the descriptive statistics). Only the above-mentioned main effect of CS-US Relation on MD scores was significant, $F(1, 120) = 11.01$, $p = .001$, $\eta^2_G = .021$. All other effects were not significant, $F(1, 120) \leq 3.49$, $ps \geq .064$, $\eta^2_G \leq .006$.

Discussion

While evidence for attitudinal ambivalence was tentative in Experiment 1, Experiment 2 more clearly supported co-occurrence information effects: a higher attitudinal ambivalence was observed when stimulus relation and co-occurrence information were supposed to diverge (contrastive condition; "prevent") than when they were supposed to converge (assimilative condition; "cause"). This effect was observed in the self-reported (although again only for the positive US valence) and mouse-tracking measures. It is also important to highlight that these effects were observed in analyses that retained only the correct trials. As a side note, we replicated previous MPT findings.

Experiment 3

In Experiments 1 and 2, we relied on Heycke and Gawronski's (2020) conditioning phase in which we used hypothetical pharmaceutical products (the CSs) that were said to cause or prevent health outcomes (the USs). We restricted some analyses to correct trials, which were trials associated with a CS evaluative classification that was in line with the evaluative implication of a given triplet (e.g., if a product is said to prevent a positive outcome, the evaluative implication is negative). These two experiments allowed us to test whether attitudinal ambivalence is higher in contrastive than in assimilative conditions, but with two important limitations. A first limitation, which has also figured in previous research (e.g., Heycke & Gawronski, 2020), is that some CS-US Relation × US Valence cells may conflict with participants' mental models. For instance, the notion that a pharmaceutical product prevents a positive health outcome may not fit well with existing representations of the pharmaceutical products' actions. Hence, CSs in the *prevents* US+ condition may be less correctly understood and retrieved from memory than CSs from the other conditions (e.g., a pharmaceutical product causing negative or positive health outcomes or a pharmaceutical product preventing a negative health outcome). As a result, it is unclear from Experiments 1 and 2 whether the results reflect higher ambivalence for CSs that prevented vs. caused health outcomes or discrepancies with mental models coming from one of the cells (possibly, the "prevent positive" one). A second limitation of Experiments 1 and 2 is that we estimated the correct encoding and retrieval of the CS-relation-US triplets meaning based on performance in the mouse-tracking task. In doing so, we possibly misclassified some CS-relation-US triplet meanings as being incorrectly encoded/retrieved. For instance, it is possible that some participants correctly remembered that a CS prevented a positive US, understood that this would imply that the CS was negative, but still found that the CS was positive.

In Experiment 3, we addressed these two shortcomings. First, we no longer used pharmaceutical products said to cause or prevent health outcomes. Instead, we adapted the relational EC procedure of Förderer and Unkelbach (2012). In their conditioning phase, Förderer and Unkelbach displayed CSs (portraits of neutral males) together with positive, neutral, or negative USs (pictures of animals or landscapes). Between exposure to each CS and US, relational information appeared that qualified the CS-US pairs: CSs were said to either "love" (assimilative relation, similar to the "cause" condition) or "loathe" (contrastive relation, similar to the "prevent" condition) the US (manipulated within participants). When CSs *love* USs, co-occurrence and relational information have the same evaluative consequences: CSs that love positive USs should be evaluated as positive, and CSs that love negative USs should be evaluated as negative. Conversely, co-occurrence and relational information should have opposite evaluative consequences when CSs *loathe* USs. Using an 8-point Likert scale, Förderer and Unkelbach found standard EC effects in the *love* condition, but reversed EC effects in the *loathe* condition. This is consistent with the notion that processing the CS-US relations is central to evaluative conditioning effects.

We decided to adapt Förderer and Unkelbach's (2012) relational EC procedure for two reasons. First, contrary to Heycke and Gawronski's (2020) procedure, it is less obvious why some CS-US Relation × US Valence cells should conflict with participants' mental models. As a result, it is a suitable procedure to address one of the two main limitations of Experiments 1 and 2. Second, varying the procedure would allow us to test whether we find evidence for higher ambivalence in contrastive vs. assimilative conditions in yet another implementation of relational information. Finding converging evidence would speak more generally to the contrastive vs. assimilative information evaluative effects rather than its specific instantiation in Heycke and Gawronski's procedure.

To address the second limitation (CS-relation-US triplets memory), we administered a role memory task at the end of the experiment (similar to studies that used a task-dissociation approach in relational evaluative conditioning procedures, see, e.g., Moran & Bar-Anan, 2013, 2016, see below for more detail).

Participants and design

The design was a 2 (US Valence: Positive vs. Negative) × 2 (CS-US Relation: Love vs. Loathe) with all factors within participants. In the conditioning phase, we counterbalanced the presentation order of the relational information (love first; loathe first) between participants.

Following the same rationale as in Experiments 1 and 2, we recruited 240 participants (our targeted sample size) on Prolific. Again, participants (1) were English speakers, (2) were right-handed, (3) had an approval rate of at least 95%, and (4) had at least 100 previous submissions. Participants did not take part in Experiments 1 and 2. Participants were paid US$2.48 for completing the study (for an estimated completion time of 15 minutes). We excluded 17 participants (four did not end the study, resulting in no data; six declared they did not pay attention to the stimuli or that they did not take their responses seriously; six had mean response times below 300 milliseconds or above 5000 milliseconds in the mouse-tracking task; we were not able to retrieve the mouse-tracking data of one participant – participants might have been excluded regarding more than one criterion). This resulted in a final sample size of 223 participants (49.33% female, one not reported; M_{age} = 37.78; SD_{age} = 14.29, one not reported).

Materials and procedure

Experiment 3 departed from Experiment 2 in five regards. First, we changed the conditioning phase and the stimuli we used. We adapted the conditioning phase from Förderer and Unkelbach (2012): CSs were said to love or loathe positive or negative animals (the USs). As CSs, we used eight nonwords used in previous studies (in their singular, not plural form): *Laapian* and *Niffian* (Ranganath & Nosek, 2008); *Belik*, *Noonnip*, *Baunif*, *Coawaq*, *Qainfin*, *Roskipp* (Navon & Bar-Anan, 2021). Contrary to Förderer and Unkelbach, we decided to use nonwords rather than portraits of males. In doing so, we reduced the risk of possible interpersonal processes related to self-other similarity (e.g., with persons as CSs, participants may like people that love [loathe] objects they like [dislike], and dislike people that love [loathe] objects they dislike [like]).

Positive and negative USs (e.g., US+: a kitten, US-: a cockroach) selected from the Open Affective Standardized Image Set (OASIS; Kurdi, Lozano, & Banaji, 2017) were repeatedly associated with the nonwords. On each trial, CS-US pairs were displayed together with relational information qualifying their link. CSs were said to either love (in one block) or loathe (in another block) the USs.

The nonwords were introduced to participants as names of alien creatures that had just arrived on Earth. Participants were asked to imagine that those aliens, to communicate, would show them their names with images of terrestrial species. Some aliens would show their names with images of animals they love, and others would show their names with images they loathe. The complete set of instructions is available at: https://osf.io/3z7bk/.

Second, each triplet was displayed ten times (vs. eight times in Experiments 1 and 2). As we used nonwords and estimated correct trials with a memory task, we wanted to increase the likelihood that participants correctly encode and retrieve the triplets. As a result, there were 80 trials total in the conditioning phase (40 in each block).

Third, we administered a memory task after the felt ambivalence task. In the memory task, we presented the CSs individually in a random order. Participants had to indicate whether each CS (1) loved a positive animal, (2) loved a negative animal, (3) loathed a positive animal, (4) loathed a negative animal, or (5) if they did not remember. Contrary to Experiments 1 and 2, we no longer estimated correct encoding and retrieval based on evaluative classifications. Instead, we used the responses to the memory task to restrict some analyses to the correct memory responses only.

Fourth, we balanced sex on Prolific to have approximately the same proportions of male and female participants. Fifth, we programmed the study with the stable version of lab.js (in Experiments 1 and 2, we used the beta version, Henninger et al., 2021).

Results

Preregistered analyses

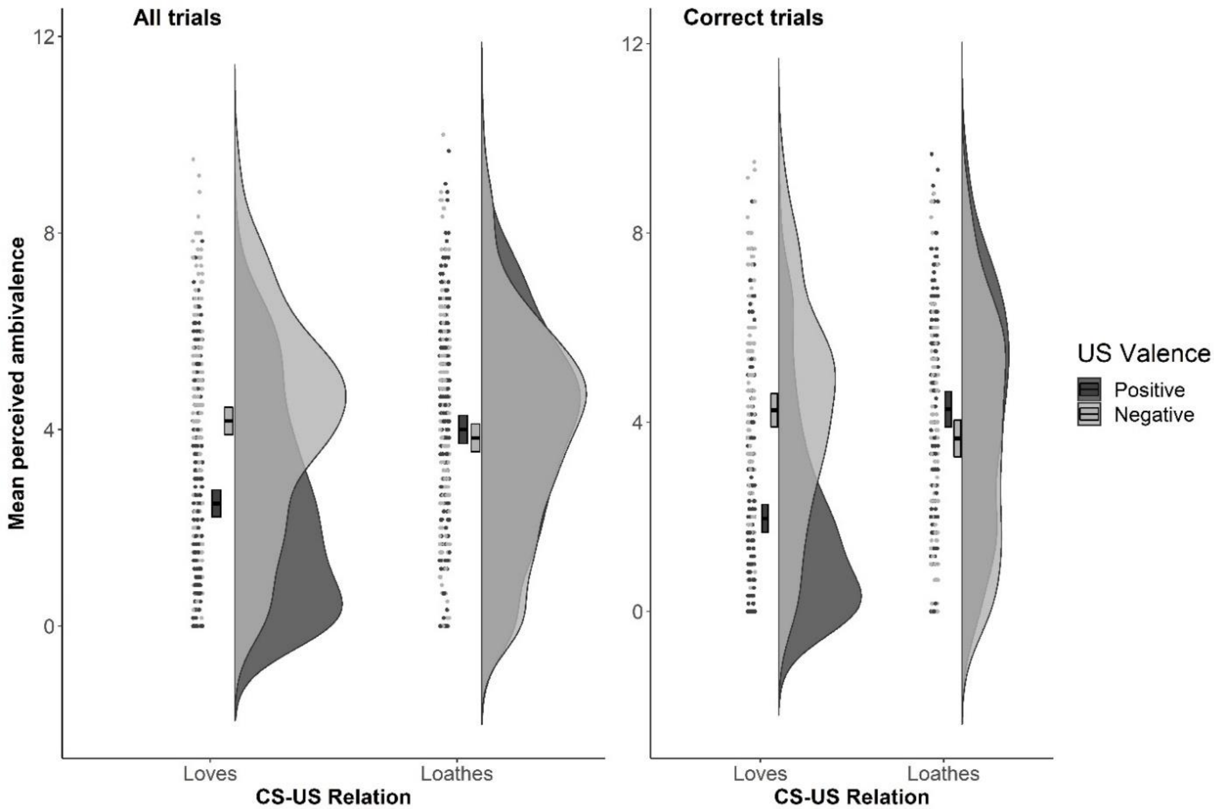

Figure 5. Mean felt ambivalence as a function of CS-US Relation and US Valence on all trials (left) and only on trials where participants provided the correct meaning-based response (right) in Experiment 3. The dots are the participants' scores (jittered). The lower and upper limits of the boxplots are the 95% confidence intervals, with the mean in between. The distributions represent the kernel probability density of the data.

All trials

Felt ambivalence scores.

In Experiments 1 and 2, we preregistered different tests for felt ambivalence (ANOVAs) and mouse-tracking measures (both paired-samples *t*-tests and ANOVAs). In Experiments 3 and 4, we preregistered the same analyses on the two measures.

We conducted a paired samples *t*-test on felt ambivalence scores with CS-US Relation as a factor. Felt ambivalence scores were higher in the Loathe than in the Love condition, $t(222) = -5.31$, $p < .001$, $d = 0.36$.

We also conducted a 2 (US Valence: Positive vs. Negative) × 2 (CS-US Relation: Love vs. Loathe) mixed ANOVA on the felt ambivalence scores[2]. The results are displayed in Figure 5 (left panel) and Table 5. The main effect of US Valence was significant, $F(1, 222) = 43.91$, $p < .001$, $\eta^2_G = .03$: Felt ambivalence scores were higher in the Negative condition than in the Positive condition. In line with the paired samples *t*-test, the main effect of CS-US Relation was significant, $F(1, 222) = 28.19$, $p < .001$, $\eta^2_G = .018$: Felt ambivalence scores were higher in the Loathe than in the Love condition. The two-way interaction between US Valence and CS-US Relation was significant too, $F(1, 222) = 60.06$, $p < .001$, $\eta^2_G = .045$. The pattern of simple effects was different from the one observed in Experiment 2 but similar to the one observed in Experiment 1. For CSs paired with positive USs, felt ambivalence scores were higher in the Loathe than in the Love condition, $t(222) = -9.15$, $p < .0001$, $d = 0.61$. In contrast, for CSs paired with negative USs, felt ambivalence scores were higher in the Love than in the Loathe condition, $t(222) = 2.16$, $p = .03$, $d = 0.15$.

Mouse-tracking measures.

Figure 6 (left panel) displays the mean mouse trajectories as a function of CS-US Relation. We conducted paired samples *t*-tests on MD and AUC scores with the CS-US Relation as a factor. MD scores were higher in the Loathe than in the Love condition, $t(222) = -3.95$, $p = .0001$, $d = 0.26$. AUC scores were also significantly higher in the Loathe than in the Love conditions, $t(222) = -2.3$, $p = .023$, $d = 0.15$.

---

[2] In the preregistered analyses, we did not include the Block' order. This is because this factor is not of primary theoretical interest, as indicated in Experiment 2.

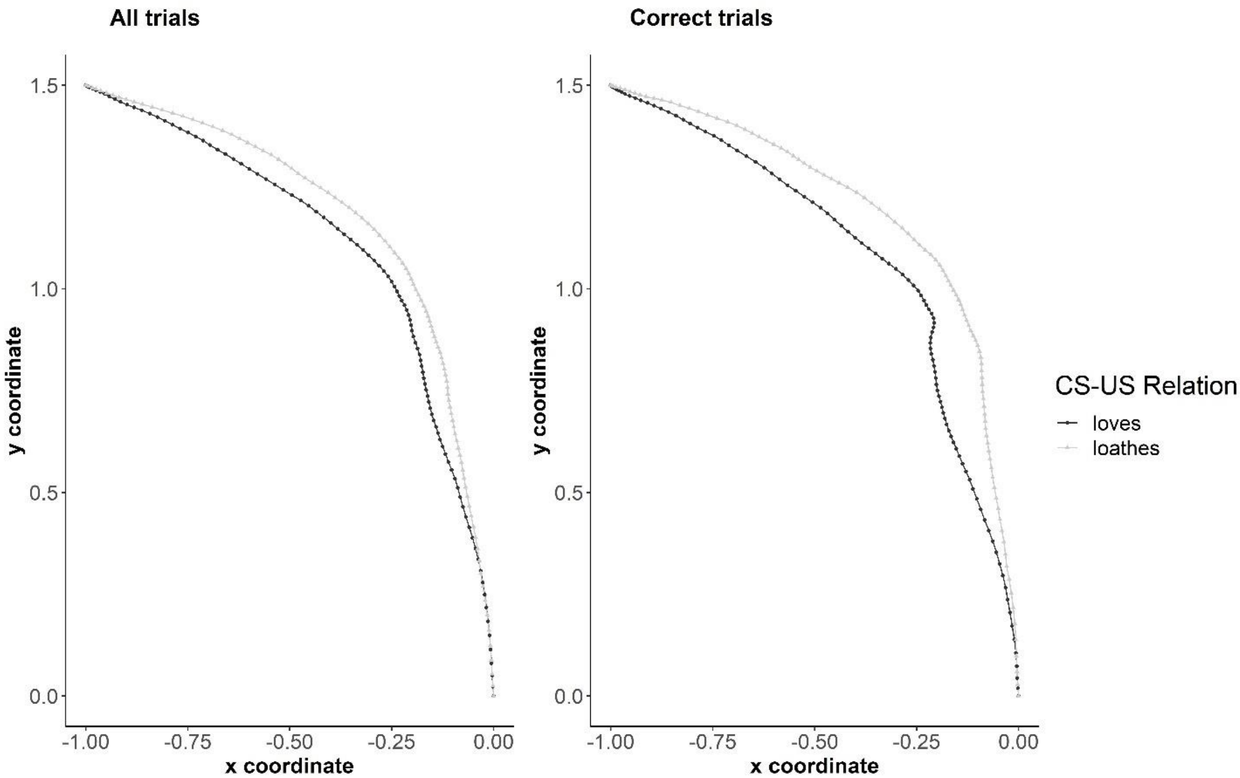

Figure 6. Average mouse trajectories as a function of CS-US Relation on each trial (left panel) and only on trials where participants provided the meaning-based correct response (right panel) in Experiment 3. All trajectories are mapped to the left.

Complementarily, we conducted a 2 (US Valence: Positive vs. Negative) × 2 (CS-US Relation: Love vs. Loathe) mixed ANOVA on MD and AUC scores (see Table 5 for the descriptive statistics). The above-mentioned main effect of CS-US Relation on MD and AUC scores were significant again, $F(1, 222) = 15.64$, $p < .001$, $\eta^2_G = .014$ (MD), $F(1, 222) = 5.27$, $p = .023$, $\eta^2_G = .005$ (AUC). We also found main effect of US Valence on both MD ($F(1, 222) = 6.42$, $p = .012$, $\eta^2_G = .005$) and AUC ($F(1, 222) = 3.91$, $p = .049$, $\eta^2_G = .004$) scores. The CS-US Relation × US Valence interactions were not significant, $Fs(1, 222) \leq 2.26$, $ps \geq .134$, $\eta^2_G \leq .002$.

RCB MPT model

Table 5. *Mean (and SD) scores in the felt ambivalence and mouse-tracking tasks as a function of US Valence and CS-US Relation in Experiment 3*

	Felt ambivalence		Mouse-tracking			
			AUC		MD	
	All	Correct	All	Correct	All	Correct
US Valence: Positive						
Love	2.49 (2.09)	1.68 (1.85)	.35 (.33)	.31 (.33)	.44 (.43)	.39 (.42)
Loathe	4.0 (2.18)	4.21 (2.49)	.42 (.34)	.45 (.4)	.59 (.47)	.63 (.52)
US Valence: Negative						
Love	4.17 (2.1)	4.18 (2.3)	.42 (.37)	.46 (.37)	.55 (.45)	.62 (.49)
Loathe	3.83 (2.15)	3.6 (2.45)	.44 (.37)	.46 (.4)	.62 (.5)	.6 (.55)

Note. AUC: Area Under the Curve; MD: Maximum Deviation.

Förderer and Unkelbach (2012)'s procedure allows us to apply Heycke and Gawronski's RCB MPT model to the evaluative classification frequencies, as the love (assimilative condition) and loathe (contrastive condition) conditions are similar to the cause and prevent conditions, respectively. The mean proportions of "positive" responses are available in Table 6.

Contrary to Experiment 2, the model did not fit the data well, $G^2(1) = 102.55$, $p < .001$ (see Table 2 for the parameter estimates).

Correct trials only

Felt ambivalence scores.

We repeated the analyses we conducted on all trials on the correct trials only (for the mean proportions of correct memory responses, see Table 6). We conducted a paired samples *t*-

Table 6. *Mean (and SD) proportions of "positive" responses in the mouse-tracking task and of correct responses in the role memory task as a function of US Valence and CS-US Relation in Experiments 3 and 4*

	Experiment 3		Experiment 4	
	"Positive"	Correct memory	"Positive"	Correct memory
US Valence: Positive				
Love	0.84 (0.27)	0.7 (0.36)	0.76 (0.32)	0.61 (0.41)
Loathe	0.48 (0.41)	0.56 (0.41)	0.41 (0.39)	0.49 (0.4)
US Valence: Negative				
Love	0.51 (0.42)	0.56 (0.41)	0.45 (0.41)	0.52 (0.42)
Loathe	0.41 (0.38)	0.49 (0.41)	0.42 (0.4)	0.47 (0.41)

test on felt ambivalence scores with CS-US Relation as a factor. Felt ambivalence scores were higher in the Loathe than in the Love condition, $t(175) = -5.94$, $p < .001$, $d = 0.45$.

We also conducted a 2 (US Valence: Positive vs. Negative) × 2 (CS-US Relation: Love vs. Loathe) mixed ANOVA (see Figure 5, right panel; see Table 4). The main effect of US Valence was significant, $F(1, 96) = 17.69$, $p < .001$, $\eta^2_G = .041$. Again, we found a significant main effect CS-US Relation, $F(1, 96) = 23.12$, $p < .001$, $\eta^2_G = .044$. The interaction between US Valence and CS-US Relation was significant, too, $F(1, 96) = 58.35$, $p < .001$, $\eta^2_G = .105$. For CSs paired with positive USs, felt ambivalence scores were higher in the Loathe than in the Love condition, $t(96) = -9.2$, $p < .0001$, $d = 0.94$. Contrary to the analyses on all trials, the difference was not significant for CSs paired with negative USs, $t(96) = 1.95$, $p = .054$, $d = 0.2$.

Mouse-tracking measures.

Figure 6 (right panel) displays the mean mouse trajectories as a function of CS-US Relation. We conducted paired samples *t*-tests on MD and AUC scores with the CS-US Relation as a factor. The effect of CS-US Relation was now significant on MD scores, $t(175) = -3.15$, $p = .002$, $d = 0.24$, but not on AUC scores, $t(175) = -1.82$, $p = .07$, $d = 0.14$.

Complementarily, we conducted a 2 (US Valence: Positive vs. Negative) × 2 (CS-US Relation: Love vs. Loathe) mixed ANOVA on MD and AUC scores (see Table 5 for the descriptive statistics). the above-mentioned main effect of CS-US Relation were significant on both MD ($F(1, 96) = 5.7$, $p = .019$, $\eta^2_G = .012$) and AUC ($F(1, 96) = 4.66$, $p = .033$, $\eta^2_G = .009$) scores. Contrary to the analyses on all trials, the main effect of US Valence was no longer significant, $F(1, 96) \leq 3.72$, $p = .057$, $\eta^2_G = .01$. The interaction between CS-US Relation and US Valence was significant on both MD ($F(1, 96) = 9.8$, $p = .002$, $\eta^2_G = .016$) and AUC ($F(1, 96) = 4.24$, $p = .042$, $\eta^2_G = .008$) scores. For CSs paired with positive USs, MD and AUC scores were higher in the Loathe than in the Love condition, $t(96) = -4.13$, $p = .0001$, $d = 0.42$ (MD); $t(96) = -3.27$, $p = .002$, $d = 0.33$ (AUC). In contrast, for CSs paired with negative USs, the simple effect of CS-US Relation was not significant, $|t|s(96) \leq 0.33$, $p \geq .744$, $ds \leq 0.03$.

Discussion

Using a different relational conditioning procedure aimed at overcoming potential limitations of Experiments 1 and 2, Experiment 3 again supports effects of co-occurrence information on ambivalence measures. On both felt and mouse-tracking measures, we observed higher ambivalence in contrastive (“loathe”) than assimilative (“love”) conditions. Of note, we also found these effects in analyses based on CSs for which participants provided correct memory responses. In addition, we again found effects involving US Valence (we will elaborate on these effects in the General Discussion). Finally, the RCB MPT model did not fit the data well.

Experiment 4

In Experiments 1 to 3, we found evidence for higher ambivalence when co-occurrence and relational information operated in opposite directions (contrastive conditions, whether “prevent” or “loathe”) than in concert (assimilative conditions, whether “cause” or “love”). However, whether the *experience* of CS-US co-occurrences is necessary to find higher ambivalence in contrastive than assimilative conditions is an open empirical question. It is possible that only being instructed about triplets would produce a similar effect. If this is the case, it would suggest that the evidence we found in Experiments 1 to 3 for co-occurrence information effects can be interpreted without reference to processes that require direct experience of the pairings (as is the case for associative processes; see the General Discussion for more on this). Conversely, if we find higher ambivalence, whether on felt or mouse-tracking measures, this would challenge the idea that direct experience is necessary to obtain co-occurrence information effects. Consequently, this finding would dovetail nicely with a propositional account of these effects. In doing so, it would challenge dual-process models that assume associative evaluative learning processes (we will further elaborate on the theoretical implications of the results in the General Discussion).

To investigate whether the direct experience of CSs and USs is necessary to find higher ambivalence in contrastive than in assimilative conditions, we adapted Experiment 3 to make it an *instructed* study. In *instructed* evaluative conditioning studies, participants are merely informed about CS-US pairings through instructions but do not directly experience the co-occurrence of CS-US pairs. The co-occurrence, if any, is only represented in a symbolic form in the wording of the instructions (see, e.g., Corneille et al., 2019; Gast & De Houwer, 2012; Hütter & De Houwer, 2017; Van Dessel et al., 2015, 2017).

Participants and design

The design was a 2 (Instructed US Valence: Positive vs. Negative) × 2 (Instructed CS-US Relation: Love vs. Loathe) with all factors within participants. In the instructed conditioning phase, we counterbalanced the presentation order of the relational information (love first; loathe first) between participants.

Following the same rationale as in Experiments 1 to 3, we recruited 240 participants (our targeted sample size) on Prolific. Again, participants (1) were English speakers, (2) were right-handed, (3) had an approval rate of at least 95%, and (4) had at least 100 previous submissions. Participants did not take part in Experiments 1 to 3. Participants were paid US $ 2.48 for completing the study (for an estimated completion time of 15 minutes). We excluded 35 participants (four did not end the study, resulting in no data; 10 declared they did not pay attention to the stimuli or that they did not take their responses seriously; 20 had mean response times below 300 milliseconds or above 5000 milliseconds in the mouse-tracking task; we were not able to retrieve the mouse-tracking data of one participant – participants might have been excluded regarding more than one criterion). This resulted in a final sample size of 205 participants (48.29% female, two responded “other”; M_{age} = 39.8; SD_{age} = 12.5).

Materials and procedure

Experiment 4 departed from Experiment 3 in three regards. First, we did not display the USs[3]. Second, we adapted the instructions of the conditioning phase so that participants were informed about the pairings in a hypothetical conditioning phase they did not take part in. The instructions, displayed on multiple screens, read as follows:

"Imagine that eight alien creatures just arrived on Earth. To communicate, some of them show us their names with images of terrestrial species they love, and others show us their names with images of terrestrial species they loathe.

In the next phase of the study, you will see the name of each alien creature together with images of animals they love or loathe. Your task is to think of the names of each alien in terms of it loving or loathing the animals displayed on the image the alien's name is paired with.

If the name of an alien is paired with the picture of a positively-valenced animal, and it says 'loves,' you should think of the alien as loving this positively-valenced animal. If the name of an alien is paired with the picture of a negatively-valenced animal, and it says 'loves,' you should think of the alien as loving this negatively-valenced animal.

If the name of an alien is paired with the picture of a positively-valenced animal, and it says 'loathes,' you should think of the alien as loathing this positively-valenced animal. If the name of an alien is paired with the picture of a negatively-valenced animal, and it says 'loathes,' you should think of the alien as loathing this negatively-valenced animal.

[3] In the preregistration of Experiment 4, we stated that we would "use" the same USs as in Experiment 3. In fact, no US were involved in the Experiment 4 as it relied on an instructed procedure (merely referring to the name of these USs), as is apparent in the study program.

In the next phase of the experiment, a given alien's name will be presented ten times together with either a positive or a negative animal. The alien will either love or loathe the animal. The positive animals are, for example, elephants hugging each other, kittens, smiling dogs, or squirrels eating. The negative animals are, for example, cockroaches, snakes, spiders, or angry ferrets. You will now see four consecutive screens.

On one screen, you will see the names of aliens that LOVE positive animals during the perception phase. On another screen, you will see the names of aliens that LOVE negative animals. Another screen will display the names of aliens that LOATHE positive animals. Another screen will display the names of aliens that LOATHE negative animals.

Each screen will be presented for 1 minute. The transition between screens will proceed automatically. Again, please think of the alien's name - image pairs in terms of the relation mentioned on the screen (loves or loathes)."

The third difference with Experiment 3 is that, for comparability with previous instructed evaluative conditioning studies (e.g., Corneille et al., 2019), we presented four screens, each displaying the two CSs allocated to each Instructed US Valence × Instructed CS-US Relation cell. The presentation time of each screen was based on the conditioning phase of our Experiment 3, in which participants saw 8 CSs displayed ten times each for 3 seconds. As a result, we displayed each CS for 60 seconds, corresponding to 10 (number of presentations in Experiment 3) × 3 (seconds) × 2 (number of CSs on each screen).

Results

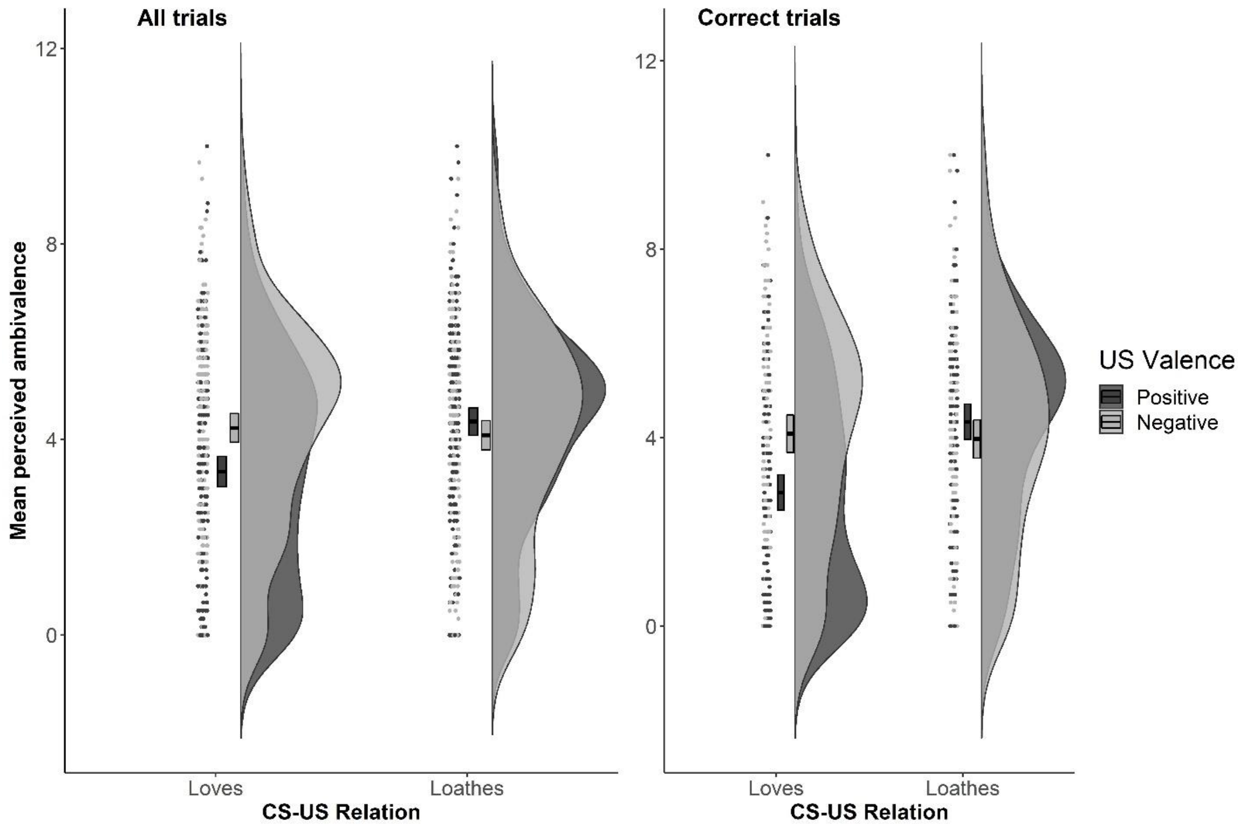

Figure 7. Mean felt ambivalence as a function of CS-US Relation and US Valence on all trials (left) and only on trials where participants provided the correct meaning-based response (right) in Experiment 4. The dots are the participants' scores (jittered). The lower and upper limits of the boxplots are the 95% confidence intervals, with the mean in between. The distributions represent the kernel probability density of the data.

Preregistered analyses

All trials

Felt ambivalence scores.

We conducted a paired samples *t*-test on felt ambivalence scores with Instructed CS-US Relation as a factor. Felt ambivalence scores were higher in the Loathe than in the Love condition, $t(204) = -3.69$, $p < .001$, $d = 0.26$.

We also conducted a 2 (Instructed US Valence: Positive vs. Negative) × 2 (Instructed CS-US Relation: Love vs. Loathe) mixed ANOVA on the felt ambivalence scores. The results are displayed in Figure 7 (left panel) and Table 7. We found a main effect of US Valence, $F(1, 204) = 9.83$, $p = .002$, $\eta^2_G = .005$: Felt ambivalence scores were higher in the Negative than in the Positive condition. In line with the paired samples *t*-test, the main effect of Instructed CS-US Relation was significant, $F(1, 204) = 13.59$, $p < .001$, $\eta^2_G = .01$: Felt ambivalence scores were higher in the Loathe condition than in the Love condition. The two-way interaction between Instructed US Valence and Instructed CS-US Relation was significant too, $F(1, 204) = 21.72$, $p < .001$, $\eta^2_G = .018$. For CSs that were said to appear with positive USs, felt ambivalence scores were higher in the Loathe than in the Love condition, $t(204) = -6.14$, $p < .0001$, $d = 0.43$. In contrast, for CSs that were said to appear with negative USs, the effect of Instructed CS-US Relation was not statistically significant, $t(204) = 0.85$, $p = .4$, $d = 0.06$.

Mouse-tracking measures.

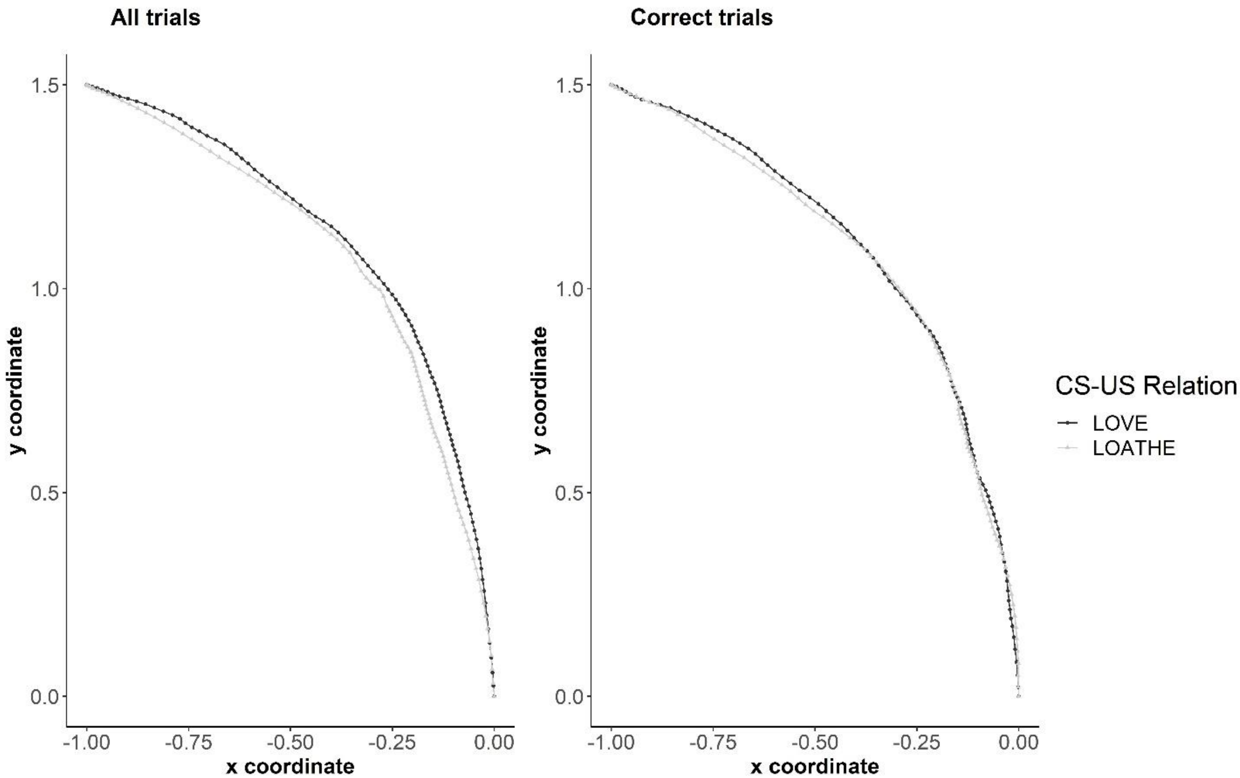

Figure 8. Average mouse trajectories as a function of CS-US Relation on each trial (left panel) and only on trials where participants provided the meaning-based correct response (right panel) in Experiment 4. All trajectories are mapped to the left.

Figure 8 (left panel) displays the mean mouse trajectories as a function of Instructed CS-US Relation. We conducted paired samples *t*-tests on MD and AUC scores with the CS-US Relation as a factor. The effect of Instructed CS-US Relation was not significant on MD ($t(204) = 1.02$, $p = .311$, $d = 0.07$) nor on AUC ($t(204) = 1.76$, $p = .08$, $d = 0.12$) scores. No main or interactive effect was significant in a 2 (Instructed US Valence: Positive vs. Negative) × 2 (Instructed CS-US Relation: Love vs. Loathe) mixed ANOVA on MD and AUC scores, $Fs(1, 205) \leq 3.14$, $ps \geq .078$, $\eta^2_G \leq .007$ (see Table 7 for the descriptive statistics).

RCB MPT model

Table 7. *Mean (and SD) scores in the felt ambivalence and mouse-tracking tasks as a function of Instructed US Valence and Instructed CS-US Relation in Experiment 4*

	Felt ambivalence		Mouse-tracking			
			AUC		MD	
	All	Correct	All	Correct	All	Correct
Instructed US Valence: Positive						
Love	3.34 (2.27)	2.03 (1.84)	.36 (.36)	.42 (.32)	.48 (.46)	.56 (.47)
Loathe	4.37 (2.04)	4.02 (2.28)	.35 (.39)	.41 (.36)	.5 (.5)	.6 (.55)
Instructed US Valence: Negative						
Love	4.24 (2.14)	3.89 (2.21)	.48 (.7)	.47 (.41)	.59 (.78)	.61 (.5)
Loathe	4.09 (2.18)	3.8 (2.31)	.41 (.63)	.48 (.46)	.51 (.54)	.61 (.55)

Note. AUC: Area Under the Curve; MD: Maximum Deviation.

As in Experiment 3, the model did not fit the data well, $G^2(1) = 63.87$, $p < .001$ (see Table 2 for the parameter estimates; see Table 6 for the mean proportions of “positive” responses).

Correct trials only

Felt ambivalence scores.

We repeated the analyses conducted on all trials on the correct trials only (for the mean proportions of correct memory responses, see Table 6). We conducted a paired samples *t*-test on felt ambivalence scores with Instructed CS-US Relation as a factor. Felt ambivalence scores were higher in the Loathe than in the Love condition, $t(148) = -4.28$, $p < .001$, $d = 0.35$.

We also conducted a 2 (Instructed US Valence: Positive vs. Negative) × 2 (Instructed CS-US Relation: Love vs. Loathe) mixed ANOVA (see Figure 7, right panel; see Table 7). The main effect of Instructed US Valence was significant, $F(1, 80) = 22.97$, $p < .001$, $\eta^2_G = .035$. Again, we found a significant main effect Instructed CS-US Relation, $F(1, 80) = 23.04$, $p < .001$, $\eta^2_G = .046$. The interaction between Instructed US Valence and Instructed CS-US Relation was also significant, $F(1, 80) = 20.09$, $p < .001$, $\eta^2_G = .055$. For CSs that were said to appear with positive USs, felt ambivalence scores were higher in the Loathe than in the Love condition, $t(80) = -6.98$, $p < .0001$, $d = 0.78$. Similar to the analyses on all trials, the difference was not significant for CSs that were said to appear with negative USs, $t(80) = .27$, $p = .79$, $d = 0.03$.

Mouse-tracking measures.

Figure 8 (right panel) displays the mean mouse trajectories as a function of Instructed CS-US Relation. We conducted paired samples *t*-tests on MD and AUC scores with the Instructed CS-US Relation as a factor. The effect of Instructed CS-US Relation was not significant whether it is on MD ($t(148) = 0.94$, $p = .347$, $d = 0.08$) or AUC ($t(148) = 1.39$, $p = .167$, $d = 0.11$) scores. No main or interactive effect was significant in a 2 (Instructed US Valence: Positive vs. Negative) × 2 (Instructed CS-US Relation: Love vs. Loathe) mixed ANOVA on MD and AUC scores, $Fs(1, 80) \leq 1.79$, $ps \geq .185$, $\eta^2_G \leq .006$ (see Table 7 for the descriptive statistics).

Discussion

We conducted Experiment 4, an “instructed” version of Experiment 3, to test whether attitudinal ambivalence is greater in contrastive than in assimilative conditions even when no direct experience of the pairings is involved. We again found greater felt ambivalence in contrastive than assimilative trials (although specifically for CSs that were said to occur with positive USs). Contrary to Experiments 1 to 3, we found no significant effect on mouse-tracking measures. As in Experiment 3, the RCB MPT model did not fit the data well. In the General Discussion, we discuss these results and their implications for theorizing attitude learning processes.

General Discussion

Whether co-occurrence information has an evaluative impact beyond relational meaning is central to attitude models and may have significant implications. Recent studies have started investigating this question by relying either on task comparison (comparing evaluative outcomes on, e.g., a direct vs. indirect evaluative task) or process dissociation procedures (estimating relational and co-occurrence parameters within a given evaluative task). In the introduction, we discussed the value of both approaches. We also argued that these procedures might be usefully complemented by the inclusion of ambivalent attitude measures. We noted that ambivalence measures allow researchers to test one consequence of the notion that processes sometimes act in concert (e.g., when a CS "causes" or "loves" a US) and sometimes act in opposition (e.g., when a CS "prevents" or "loathes" a US). That is, when co-occurrence information and stimulus relation support opposite evaluations, attitudinal ambivalence should be higher than when they support the same evaluation. This hypothesis cannot be tested with task comparison or process dissociation studies. Of critical interest too, ambivalence measures allow researchers to carry out this test within a single evaluative task (which cannot be done in task comparison studies) and to restrict analyzes to correct trials only, for which there is a better chance of correct encoding and retrieval of the relational meaning of the information (which cannot be done in process dissociation studies).

In four preregistered experiments, we found support for higher ambivalence when co-occurrence information and stimulus relation have divergent evaluative implications (contrastive conditions) than when they have convergent evaluative implications (assimilative conditions). Specifically, we systematically found higher self-reported ambivalence in contrastive than assimilative conditions, although this effect was restricted to CSs paired with positive USs. In the case of mouse-tracking measures, we found higher ambivalence in contrastive conditions in the three experiential experiments (Experiments 1 to 3) in all analyses that were restricted to correct trials. Mouse-tracking measures were not sensitive to the manipulations in Experiment 4, which implemented an instructed conditioning phase (i.e., when no direct experience of the CS-US pairings was involved).

While there are some limitations to our results (see below), they support, on the whole, the conclusion that co-occurrence information influences CS evaluations even when the relational meaning is correctly encoded and retrieved. Importantly, this conclusion is supported by two relational procedures (cause/prevent in Experiments 1 and 2; love/loathe in Experiments 3 and 4). The latter procedure included both an experiential (Experiment 3) and an instruction-based conditioning phase (Experiment 4).

In the remainder of this General Discussion, we elaborate on how single and dual-process models of attitude learning can account for the present results. We also point out limitations to the current research and suggest how ambivalence measures may help advance our understanding of attitude acquisition in future studies.

Interpretation of the present findings

To interpret the present findings, we need a typology of the processing stages involved in their production. In the present experiments, all we know is how participants evaluated the stimuli and felt about them. These outcomes are the end product of multiple processes, including, but not limited to, attention paid to the information, how the information was interpreted, the consolidation of this interpretation, its subsequent retrieval, and how the recollected memories were used to produce an evaluation (see, e.g., Gast, 2018; Hütter & Rothermund, 2021; Stahl & Aust, 2018). Because the current findings are, for a part, based on correct responses, this weakens the possibility that the results were driven by trials in which participants were not attentive to or did not correctly interpret, retrieve, or use the relational information. The findings suggest that people may acquire an attitudinal representation whose structure is characterized by positive and negative components in the contrastive (i.e., prevent/loathe) conditions and by either two positive or two negative components in the assimilative (i.e., cause/love) conditions.

We now discuss how dual- and single-process models account for the co-occurrence information effect we found in the form of higher ambivalence in contrastive conditions as opposed to assimilative ones. According to dual-process theories of attitude learning (e.g., the Associative-Propositional Evaluation model, APE; Gawronski & Bodenhausen, 2006, 2011, 2014, 2018; the Systems of Evaluation Model, SEM; Rydell & McConnell, 2006; McConnell & Rydell, 2014), evaluations result from both associative and propositional learning processes. A critical operating principle of associative processes is that they automatically register mere co-occurrences between stimuli. As a result, associative processes are commonly assumed to be insensitive to the truth value and the meaning of the pairs. Conversely, propositional processes are thought to capture the meaning and truth value of the pairs. For dual-process models, the existence of two independent (convergent or divergent) evaluations in relational EC procedures is straightforward: a co-occurrence-based evaluation is produced by associative processes (which only register unqualified associations between stimuli), while a meaning-based evaluation is produced by propositional processes.

Consistent with dual-process models, we found that co-occurrence information influences participants' evaluations in addition to stimulus relation. However, the fact that felt ambivalence measures were higher in a contrastive condition than in an assimilative one in an instructed procedure (Experiment 4), challenges views that imply the contribution of associative processes. That we found more deviated mouse trajectories in contrastive conditions in experiential but not instructed procedures may be interpreted as selectively supporting associative learning processes. Associative models have long assumed that indirect tasks are more sensitive to associative processes than direct and more deliberate measures. Hence, one may argue that, as associative processes can no longer apply in instructed procedures (as no direct experience of the CS-US pairings is involved), co-occurrence effects would not be found in the mouse-tracking task of an instruction-based study like Experiment 4. However, we think that this interpretation is misleading. This is because evaluations were fully deliberate in the mouse-tracking tasks: participants were asked to evaluate the CSs and did so without time pressure (although participants were asked to move their pointing device quickly). Furthermore, recent research has supported the sensitivity of indirect tasks, such as the Implicit Association Test, to relational information (Bading et al., 2020). More generally, the assumption that indirect and direct tasks are differently sensitive to propositional and associative processes has now been widely questioned (see, e.g., Corneille & Hütter, 2020; De Houwer, 2009; Gawronski, 2019; for an extension to physiological measures, see Corneille & Mertens, 2020).

Turning to single-process models, we note that, as was the case of dual-process models, they account for a portion of the results. Single-propositional models reject the existence of associative learning processes and only posit the existence of propositional processes (e.g., the Integrative Propositional Model, IPM; De Houwer, 2018). Propositional models assume that the evaluative information is encoded in the form of symbolic propositions that bear relational meaning. For the IPM, co-occurrence can influence evaluations even when only one proposition containing the correct stimulus relation has been encoded, as incomplete information about the triplets can be retrieved (the partial retrieval hypothesis). For instance, participants may sometimes first remember that a CS is somehow related to a US and only then retrieve the information that this CS caused or prevented the US. However, this account may not best explain the presence of ambivalence on the felt ambivalence measures, as these measures suggest the experience of positive and negative evaluations that occur at nearly the same time, if not simultaneously. According to the partial retrieval hypothesis, a participant performing the felt ambivalence task may first partially remember that a CS co-occurred with a positive US and only then retrieve the full proposition that the CS was said to prevent the US. If this were the case, there would be no reason for them to feel conflicted, indecisive, or mixed about the CS. As soon as the full proposition is retrieved, it should not generate ambivalence. Moreover, such an account does not explain why we found higher ambivalence in contrastive than assimilative conditions on felt ambivalence but not on mouse-tracking measures in the instruction-based study (Experiment 4).

From a propositional perspective, another possibility is that two propositions are encoded and retrieved from memory: one that captures the meaning of the information (e.g., X is bad for health) and one that captures the co-occurrence information (e.g., X reminds me of cancer). This would be consistent with the results we obtained here. Finally, a single-process episodic memory model (Gast, 2018; Stahl & Aust, 2018; Stahl & Heycke, 2016) would be very similar to the rationale applied to the single-propositional model. From an episodic memory perspective, various (possibly inconsistent) pieces of information are retrieved from memory with varying degrees of confidence, and their integration underpins evaluations.

Limitations

A first caveat of the present research is that we estimated attitudinal ambivalence at the evaluation production stage, but we have drawn tentative conclusions about attitude acquisition (not retrieval) processes. As we only included analyses of correct trials, we have some guarantee that the evidence for attitudinal ambivalence holds when stimulus relation and US valence are correctly encoded and retrieved. However, the experimental designs we used do not offer direct evidence for the claim that learning processes are responsible for the observed attitudinal ambivalence. This is only one plausible explanation that can be drawn from our findings. Larger response competition (observed in the mouse-tracking task) in the assimilative rather than the contrastive conditions could result from lower confidence in responses, for example. That is, when co-occurrence information and stimulus relation have divergent evaluative implications, participants may find it more difficult to understand or retrieve the correct meaning because they need to integrate two sources of information to infer an evaluation (while such inference is not necessary when co-occurrence and stimulus relations have convergent implications). This explanation might apply to our four experiments, whether they implemented cause/prevent (Experiments 1 & 2) or love/loathe (Experiments 3 & 4) procedures. Future research could test whether a conditioning phase (instruction-based or experienced) is required at all for attitudinal ambivalence to occur. One way to address this question could consist of administering only an evaluation stage where CSs are said to cause or prevent either positive or negative USs. If evidence for attitudinal ambivalence were to be found in such studies, this would suggest that a response-stage-based account suffices to explain the current results (as they would be observed in the absence of a prior conditioning phase, whether experienced or only instructed). Such a line of research may also help determine whether inferences are more difficult to make and retrieve when CSs and USs have contrastive rather than assimilative relations.

Another possibility is that contrastive CS-US relations may induce less extreme CS evaluations than assimilative relations (see Bading et al., 2022). This is because contrastive relations may conflict more with mental models. Although such an account may well apply to Experiments 1 and 2, which featured potential mental model issues (see the introduction of Experiment 3 above), it is unclear why it would apply to Experiments 3 and 4, which were designed to alleviate this problem. Second, and even more importantly, this neutrality account fails to explain the higher felt ambivalence observed in contrastive rather than assimilative conditions. We do not think it would be parsimonious to explain the mouse-tracking results in terms of higher neutrality in the contrastive than assimilative condition while explaining the similar results on the felt ambivalence measures in terms of higher ambivalence. Finally, it is worthwhile repeating that the mouse-tracking task has been validated as a measure of ambivalence (Schneider et al., 2015) rather than as a measure of neutrality (Schneider & Mattes, 2021).

A second general limitation to consider is that the level of ambivalence was generally moderate in our four experiments. Even if we found higher ambivalence in contrastive than assimilative conditions, the effect sizes were rather small overall. The presence of non-zero ambivalence in the assimilative condition, however, is interesting in and of itself[4]. It raises the question of why the CSs partly acquired a negative (positive) valence after being paired with both positive (negative) co-occurrence and positive (negative) relational information. A possibility is that the stimuli were already partly ambivalent (rather than purely neutral) before their conditioning. This possibility cannot be excluded, especially when considering that the CSs were pharmaceutical products in Experiments 1 and 2. As we used nonwords in Experiments 3 and 4, this is less of a concern in these studies. Although this does not represent a confounding factor in the present experiments, future research may pre-test CS ambivalence over and above CS neutrality (e.g., Schneider et al., 2017, 2021).

More generally, the overall low level of ambivalence we observed may suggest that co-occurrence and relational information had weak influences or that one influence largely dominated the other. The MPT modeling delivered relatively strong parameter estimates in Experiments 1 and 2 (e.g., in Experiment 2, we found $R = .4$ and $C = .38$, both significantly above 0). This may indicate that both processes operated to a large extent (although C is conditional upon the absence of R; meaning that, in Experiment 2, relational meaning drove responses in 40% of the trials, and co-occurrence drove responses in 23% of the trials $[1\text{-}R]*C =$

---

[4] In the present research, we were mainly interested in whether attitudinal ambivalence was larger when co-occurrence information and stimulus relation had divergent rather than convergent implications. Attitudinal ambivalence could also be estimated in each condition by comparing scores (MD and AUC in the mouse-tracking task; felt ambivalence on the scales) to 0. In non-preregistered analyses of the four experiments, we systematically found that all ambivalence scores were significantly above 0.

.23). As we have explained, however, one can hardly conclude from MPT parameters that these influences operated on the same evaluations (as any given evaluation entered the relational meaning branch *or* the co-occurrence branch in the absence of the relational meaning information's influence). It is difficult to conclude whether both processes had low influence or whether one was dominant from the ambivalence measures. If the latter, however, one should conclude that relational meaning dominated co-occurrence information. This is because ambivalence was best observed in correct CS evaluations, which by definition captured the relational meaning of the information (whether in the evaluative classification task in Experiments 1 and 2 or a subsequent memory task in Experiments 3 and 4).

A third limitation is that there were several findings we cannot yet easily account for. The pattern of results on the two measures of trajectories' curvature in the mouse-tracking task – namely maximum deviation and area under the curve – often diverged according to their sensitivity to the experimental manipulations. This was especially the case in Experiments 1 and 2 with the cause/prevent procedure and less so in Experiment 3 with the love/loathe procedure. Maximum deviation and area under the curve were higher when co-occurrence information and stimulus relation had divergent rather than convergent implications (only on correct trials and only for maximum deviation scores in Experiment 1; both overall and on correct trials in Experiments 2 and 3), but this was not the case in all statistical tests. Given that maximum deviation and area under the curve were highly correlated (Experiment 1: r = .91; Experiment 2: r = .86; Experiment 3: r = .93; Experiment 4: r = .88, as is typically the case, e.g., Cummins & De Houwer, 2020; Schneider et al., 2015) and they demonstrated overall the same patterns of results, we cannot offer a straightforward explanation or speculate as to why they sometimes diverged in their sensitivity to the cause/prevent procedure.

Concerning the felt ambivalence task, we found a significant interaction between US valence and stimulus relation in the four experiments. Felt ambivalence was only higher in the prevent/loathe conditions than cause/love conditions for positive USs (in the four experiments). By contrast, felt ambivalence was *lower* in the prevent/loathe conditions in Experiments 1 and 3 but not statistically different from the cause/love conditions in Experiments 2 and 4 (this was also the case on the maximum deviation for correct trials only in Experiment 1). This result may be important as it suggests that co-occurrence information and stimulus relation could be two related, not orthogonal, sources of evaluative information, contrary to what has been hypothesized in MPT models. One possibility is that the evaluative inferences one has to make are more complex when negative rather than positive USs are involved[5]. For instance, in the love/loathe procedure (Experiments 3 & 4), we used valenced words to create assimilative (“love,” which is positive) and contrastive (“loathe,” which is negative) CS-US relations. The evaluative inference may be straightforward when a CS is said to love a positive US (i.e., a positive inference). When a CS is said to loathe a negative US, the evaluation may be more complex because participants have to infer a positive evaluation based on two negative pieces of information. When a CS is said to love a negative US, and when a CS is said to loathe a positive US, the mix of positive and negative valences may confuse participants. As a result, negative USs would always be associated with complex inferences (whether the CS-US relation is assimilative or contrastive), while positive USs would only be associated with complex inferences in the contrastive condition. Given that higher inference complexity may make

---

[5] We thank one of the anonymous reviewers for raising this possibility.

participants confused or ambivalent, having negative USs always involved in complex inferences may make them less suited to detect the effect of relational meaning on ambivalence measures than positive USs. Although we acknowledge that this explanation is entirely ad-hoc and speculative, future research could address the issue of inference complexity in order to better understand the specific moderators involved in relational evaluative conditioning procedures.

In Experiment 4, which implemented an instruction-based conditioning phase, we found higher felt ambivalence in the contrastive than in the assimilative condition but no difference in mouse-tracking measures. The fact that the felt ambivalence was sensitive to our instruction manipulation but not the mouse-tracking measures is not readily accounted for by either dual-process or propositional single-process models. This is because these models do not predict different effects on felt ambivalence and mouse-tracking measures. Given that CS-US pairings are not directly experienced in instruction-based procedures, co-occurrence information effects are more likely to be based on propositional than associative processes. As a result, finding higher felt ambivalence in contrastive as opposed to assimilative (instructed) conditions but not on mouse-tracking measures is problematic for dual-process models. Conversely, single-process models have issues explaining why we did not find effects on the mouse-tracking task as co-occurrence information effects are supposed to be based on propositional processes both in the felt ambivalence and mouse-tracking tasks.

At the moment, we cannot offer a complete account as to why the instruction-based procedure made mouse-tracking measures insensitive to our manipulations but not felt ambivalence measures. One possibility[6] is that relational information and US valence are more difficult to retrieve in the instruction-based study we conducted than in experiential studies, possibly because only verbal information was available (as opposed to pictures in experiential studies). As displayed in Table 6, CS role memory performance was descriptively lower in Experiment 4 (instruction-based) than in Experiment 3 (experiential), although they were overall similar. If mouse-tracking measures capture quicker evaluations than felt ambivalence measures, participants may often have responded to the mouse-tracking task without retrieving their evaluative inferences in the instruction-based study (Experiment 4).

Future directions

It is important to point out that the influence of co-occurrence information examined in this article speaks to operating principles of evaluative learning (i.e., what the evaluative learning processes do). Operating principles are conceptually different from operating conditions. The latter is concerned with when evaluative learning emerges and, more specifically, whether evaluative learning occurs under automaticity conditions (e.g., unawareness, efficiency, uncontrollability, goal-independence; see, e.g., Gawronski & Bodenhausen, 2009, 2014; Corneille & Stahl, 2018). A particularly interesting question is how the influence of co-occurrence information depends on encoding conditions. For instance, Heycke and Gawronski (2020, Experiments 2 and 3) manipulated encoding conditions in a relational EC procedure. Regarding the "efficiency" issue, they found that limiting processing time at encoding decreased the RCB MPT model's *R* parameter without affecting the *C* parameter. Conducting similar

[6] We thank one of the anonymous reviewers for raising this possibility.

research but using attitudinal ambivalence measures instead may help clarify the contribution of acquisition-related processes to the current set of findings.

Another worthwhile avenue for future research is to use ambivalence measures to investigate another operating condition: (un)controllability. It has been proposed that CSs uncontrollably acquire the valence of their paired US in evaluative conditioning procedures. That is, participants cannot resist the evaluative influence of the US pairing even when trying to avoid it (e.g., Hütter & Sweldens, 2018; Gawronski & Brannon, 2021; Gawronski et al., 2015; Peters & Gawronski, 2011; Waroquier et al., 2022). In two studies, Hütter and Sweldens supported this conclusion by relying on a process dissociation procedure that found evidence for the contribution of both controlled and uncontrolled attitude formation processes to EC effects. Specifically, two above-0 parameter estimates were observed in their studies: one that captured the controlled application of the instruction to form standard or reversed impressions of the CSs, and one that was interpreted as reflecting the direct and uncontrollable transfer of the US valence to the CS. Here too, the use of ambivalence measures may help clarify whether convergent processes act in the standard instruction condition (i.e., instruction to apply the US valence to the CS) and divergent processes operate in the reversal instruction condition (i.e., instruction to apply the opposite US valence to the CS; for preliminary findings, see Corneille et al., 2019). In other words, ambivalence should be higher in the reversal than the standard instruction condition.

Conclusion

By using attitudinal ambivalence measures, the present investigation provides new and significant insights into the role of co-occurrence information in evaluative processes. It indicates that co-occurrence information influences evaluations *even when* there is an indication that the relational information is correctly encoded and retrieved. It also suggests that co-occurrence information and relational meaning can influence evaluations closely in time. Both single-propositional and dual-process attitude learning models face challenges in explaining the present results. Future studies may examine whether the enhanced ambivalence observed in the assimilative condition is automatically (e.g., uncontrollably, efficiently) established; examine the contribution of learning and post-learning processes to the effects; and further investigate the behavioral implications of the ambivalence created when the evaluative influences of relational meaning and co-occurrence information diverge.

Context of the Research

For the last 15 years, our laboratory has investigated whether evaluative learning occurs under conditions of automaticity. Specifically, we tested whether evaluative learning emerges without CS-US awareness, uncontrollably, efficiently, and independently of goals. Overall, we found little evidence for automatic evaluative learning. These questions pertain to the operating conditions of evaluative learning: *when* it emerges. With the current research, we turned to the study of operating principles: *what* the processes do. The present “ambivalence-based” approach supports that evaluative learning can be based on co-occurrence information. It remains unclear, however, whether this result is best explained in terms of single- (propositional) or dual- (associative and propositional) process models (or other models). In future research, we intend to use ambivalence measures to continue testing operating principles of evaluative learning and provide new tests of its operating conditions.

References

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., van Langen, J., & Kievit, R. A. (2021). Raincloud plots: a multi-platform tool for robust data visualization. *Wellcome Open Research, 4*, 63. https://doi.org/10.12688/wellcomeopenres.15191.2

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*(1), 128–143. https://doi.org/10.1080/02699931.2019.1604322

Bading, K. C., Barth, M., Rothermund, K., Aust, F., & Stahl, C. (2022). *Does mere co-occurrence affect evaluation? MPT modelling of relational evaluative conditioning cannot (yet) tell*. [Preprint.] *PsychArchives*. https://doi.org/10.23668/psycharchives.5381

Ben-Shachar, M., Lüdecke, D., & Makowski, D. (2020). effectsize: Estimation of Effect Size Indices and Standardized Parameters. *Journal of Open Source Software*, *5*(56), 2815. https://doi.org/10.21105/joss.02815

Calanchini, J. (2020). How Multinomial Processing Trees Have Advanced, and Can Continue to Advance, Research Using Implicit Measures. *Social Cognition*, *38*(Supplement), s165–s186. https://doi.org/10.1521/soco.2020.38.supp.s165

Calanchini, J., Rivers, A. M., Klauer, K. C., & Sherman, J. W. (2018). Multinomial processing trees as theoretical bridges between cognitive and social psychology. *Psychology of Learning and Motivation*, 39–65. https://doi.org/10.1016/bs.plm.2018.09.002

Conner, M., & Armitage, C. J. (2008). Attitudinal ambivalence. In W. D. Crano & R. Prislin (Eds.), *Attitudes and attitude change* (pp. 261–286). Psychology Press.

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. https://doi.org/10.1016/j.jesp.2019.103841

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*(3), 212–232. https://doi.org/10.1177/1088868320911325

Corneille, O., & Mertens, G. (2020). Behavioral and Physiological Evidence Challenges the Automatic Acquisition of Evaluations. *Current Directions in Psychological Science*, *29*(6), 569–574. https://doi.org/10.1177/0963721420964111

Corneille, O., & Stahl, C. (2018). Associative Attitude Learning: A Closer Look at Evidence and How It Relates to Attitude Models. *Personality and Social Psychology Review*, *23*(2), 161–189. https://doi.org/10.1177/1088868318763261

Cummins, J., & De Houwer, J. (2020). The Shape of Belief: Developing a Mousetracking-Based Relational Implicit Measure. *Social Psychological and Personality Science*, 194855062097801. https://doi.org/10.1177/1948550620978019

De Houwer, J. (2007). A Conceptual and Theoretical Analysis of Evaluative Conditioning. *The Spanish Journal of Psychology*, *10*(2), 230–241. https://doi.org/10.1017/s1138741600006491

De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, *37*(1), 1–20. https://doi.org/10.3758/lb.37.1.1

De Houwer, J. (2018). Propositional Models of Evaluative Conditioning. *Social Psychological Bulletin*, *13*(3), 1-21. https://doi.org/10.5964/spb.v13i3.28046

Erdfelder, E., Auer, T.-S., Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial Processing Tree Models. *Journal of Psychology*, *217*(3), 108–124. https://doi.org/10.1027/0044-3409.217.3.108

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Förderer, S., & Unkelbach, C. (2012). Hating the cute kitten or loving the aggressive pit-bull: EC effects depend on CS–US relations. *Cognition & Emotion*, *26*(3), 534–540. https://doi.org/10.1080/02699931.2011.588687

Gast, A. (2018). A Declarative Memory Model of Evaluative Conditioning. *Social Psychological Bulletin*, *13*(3), e28590. https://doi.org/10.5964/spb.v13i3.28590

Gast, A., & De Houwer, J. (2012). Evaluative Conditioning without Directly Experienced Pairings of the Conditioned and the Unconditioned Stimuli. *Quarterly Journal of Experimental Psychology*, *65*(9), 1657–1674. https://doi.org/10.1080/17470218.2012.665061

Gawronski, B. (2019). Six Lessons for a Cogent Science of Implicit Bias and Its Criticism. *Perspectives on Psychological Science*, *14*(4), 574–595. https://doi.org/10.1177/1745691619826015

Gawronski, B., Balas, R., & Creighton, L. A. (2013). Can the Formation of Conditioned Attitudes Be Intentionally Controlled? *Personality and Social Psychology Bulletin*, *40*(4), 419–432. https://doi.org/10.1177/0146167213513907

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*(5), 692–731. https://doi.org/10.1037/0033-2909.132.5.692

Gawronski, B., & Bodenhausen, G. V. (2009). Operating principles versus operating conditions in the distinction between associative and propositional processes. *Behavioral and Brain Sciences, 32*(2), 207–208. https://doi.org/10.1017/s0140525x09000958

Gawronski, B., & Bodenhausen, G. V. (2011). The Associative–Propositional Evaluation Model. *Advances in Experimental Social Psychology*, 59–127. https://doi.org/10.1016/b978-0-12-385522-0.00002-0

Gawronski, B., & Bodenhausen, G. V. (2014). Implicit and Explicit Evaluation: A Brief Review of the Associative-Propositional Evaluation Model. *Social and Personality Psychology Compass, 8*(8), 448–462. https://doi.org/10.1111/spc3.12124

Gawronski, B., & Bodenhausen, G. V. (2014). The associative-propositional evaluation model: Operating principles and operating conditions of evaluation. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 188-203). Guilford Press.

Gawronski, B., & Bodenhausen, G. V. (2018). Evaluative Conditioning From the Perspective of the Associative-Propositional Evaluation Model. *Social Psychological Bulletin, 13*(3). https://doi.org/10.5964/spb.v13i3.28024

Gawronski, B., Luke, D. M., & Ng, N. L. (2021). Is less of an unhealthy ingredient healthy or unhealthy? Effects of mere co-occurrence and quantitative relations on attribute judgments. *Journal of Experimental Social Psychology*, 104193. https://doi.org/10.1016/j.jesp.2021.104193

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*(5), 556–568. https://doi.org/10.1037/emo0000078

Gawronski, B., & Brannon, S. M. (2021). Attitudinal Effects of Stimulus Co-Occurrence and Stimulus Relations: Range and Limits of Intentional Control. *Personality and Social Psychology Bulletin*, 014616722098290. https://doi.org/10.1177/0146167220982906

Glaser, T., & Walther, E. (2012). One but not the same: Evaluative conditioning with mixed-valence USs. Learning and Motivation, 43(3), 144–154. https://doi.org/10.1016/j.lmot.2012.06.005

Glaser, T., Woud, M. L., Labib Iskander, M., Schmalenstroth, V., & Vo, T. M. (2018). Positive, negative, or all relative? Evaluative conditioning of ambivalence. *Acta Psychologica, 185*, 155–165. https://doi.org/10.1016/j.actpsy.2018.02.006

Henninger, F., & Kieslich, P. J. (2020). *Mousetrap-Web: Mouse-tracking for the browser*. Manuscript submitted for publication. Retrieved at: https://github.com/FelixHenninger/mousetrap-web/tree/master/manuscript

Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J., & Hilbig, B. E. (2021). lab.js: A free, open, online study builder. *Behavior Research Methods*, *54*, 556-573. https://doi.org/10.3758/s13428-019-01283-5

Heycke, T., & Gawronski, B. (2020). Co-occurrence and relational information in evaluative learning: A multinomial modeling approach. *Journal of Experimental Psychology: General*, *149*(1), 104–124. https://doi.org/10.1037/xge0000620

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin*, *136*(3), 390–421. https://doi.org/10.1037/a0018916

Hu, X., Gawronski, B., & Balas, R. (2016). Propositional Versus Dual-Process Accounts of Evaluative Conditioning. *Personality and Social Psychology Bulletin*, *43*(1), 17–32. https://doi.org/10.1177/0146167216673351

Hütter, M., & De Houwer, J. (2017). Examining the contributions of memory-dependent and memory-independent components to evaluative conditioning via instructions. *Journal of Experimental Social Psychology*, *71*, 49–58. https://doi.org/10.1016/j.jesp.2017.02.007

Hütter, M., & Klauer, K. C. (2016). Applying processing trees in social psychology. *European Review of Social Psychology*, 27(1), 116–159. https://doi.org/10.1080/10463283.2016.1212966

Hütter, M., & Rothermund, K. (2020). Automatic processes in evaluative learning. *Cognition and Emotion*, *34*(1), 1–20. https://doi.org/10.1080/02699931.2019.1709315

Hütter, M., & Sweldens, S. (2018). Dissociating Controllable and Uncontrollable Effects of Affective Stimuli on Attitudes and Consumption. *Journal of Consumer Research*, *45*(2), 320–349. https://doi.org/10.1093/jcr/ucx124

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513–541. https://doi.org/10.1016/0749-596x(91)90025-f

Jin, R., Xia, T., Gawronski, B., & Hu, X. (2022). Attitudinal Effects of Stimulus Co-occurrence and Stimulus Relations: Sleep Supports Propositional Learning Via Memory

Consolidation. *Social Psychological and Personality Science*, 194855062110676. https://doi.org/10.1177/19485506211067673

Kassambara, A. (2020). *ggpubr: 'ggplot2' Based Publication Ready Plots.* R package version 0.4.0. Retrieved from https://CRAN.R-project.org/package=ggpubr

Kieslich, P. J., & Henninger, F. (2017). Mousetrap: An integrated, open-source mouse-tracking package. *Behavior Research Methods*, *49*(5), 1652–1667. https://doi.org/10.3758/s13428-017-0900-z

Kieslich, P. J., Henninger, F., Wulff, D. U., Haslbeck, J. M. B., & Schulte-Mecklenbeck, M. (2019). Mouse-Tracking: A Practical Guide to Implementation and Analysis. In *A Handbook of Process Tracing Methods* (pp. 111–130). Routledge.

Kukken, N., Hütter, M., & Holland, R. W. (2019). Are there two independent evaluative conditioning effects in relational paradigms? Dissociating the effects of CS-US pairings and their meaning. *Cognition and Emotion*, *34*(1), 170–187. https://doi.org/10.1080/02699931.2019.1617112

Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the Open Affective Standardized Image Set (OASIS). *Behavior Research Methods*, *49*(2), 457–470. https://doi.org/10.3758/s13428-016-0715-3

Lange, K., Kühn, S., & Filevich, E. (2015). Correction: “Just Another Tool for Online Studies” (JATOS): An Easy Solution for Setup and Management of Web Servers Supporting Online Studies. *PLOS ONE*, *10*(7), e0134073. https://doi.org/10.1371/journal.pone.0134073

Lenth, R. (2020). emmeans: Estimate Marginal Means, aka Least-Squares Means. R package version 1.5.2-1. https://CRAN.R-project.org/package=emmeans

McConnell, A. R., & Rydell, R. J. (2014). The systems of evaluation model: A dual-systems approach to attitudes. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual process theories of the social mind* (pp. 204-217). Guilford.

Moran, T., & Bar-Anan, Y. (2013). The effect of object–valence relations on automatic evaluation. *Cognition & Emotion, 27*(4), 743–752. https://doi.org/10.1080/02699931.2012.732040

Moran, T., Bar-Anan, Y., & Nosek, B. A. (2015). Processing goals moderate the effect of co-occurrence on automatic evaluation. *Journal of Experimental Social Psychology, 60,* 157–162. https://doi.org/10.1016/j.jesp.2015.05.009

Moran, T., Bar-Anan, Y., & Nosek, B. A. (2016). The assimilative effect of co-occurrence on evaluation above and beyond the effect of relational qualifiers. *Social Cognition*, *34*(5), 435-461. https://doi.org/10.1521/soco.2016.34.5.435

Moran, T., & Bar-Anan, Y. (2019). The effect of co-occurrence and relational information on speeded evaluation. *Cognition and Emotion*, *34*(1), 144–155. https://doi.org/10.1080/02699931.2019.1604321

Navon, M., & Bar-Anan, Y. (2021). When Does Group Information Influence the Automatic and Deliberate Judgment of Individual Group Members? Preprint. https://doi.org/10.31234/osf.io/zpuf9

Payne, B. K., & Bishara, A. J. (2009). An integrative review of process dissociation and related models in social cognition. *European Review of Social Psychology*, *20*, 272-314. https://doi.org/10.1080/10463280903162177

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*(4), 557–569. https://doi.org/10.1177/0146167211400423

Priester, J. R., & Petty, R. E. (1996). The gradual threshold model of ambivalence: Relating the positive and negative bases of attitudes to subjective ambivalence. *Journal of Personality and Social Psychology*, *71*(3), 431–449. https://doi.org/10.1037/0022-3514.71.3.431

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review*, *95*(3), 318–339. https://doi.org/10.1037/0033-295x.95.3.318

Rothman, N. B., Pratt, M. G., Rees, L., & Vogus, T. J. (2017). Understanding the Dual Nature of Ambivalence: Why and When Ambivalence Leads to Good and Bad Outcomes. *Academy of Management Annals*, *11*(1), 33–72. https://doi.org/10.5465/annals.2014.0066

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology*, *91*(6), 995–1008. https://doi.org/10.1037/0022-3514.91.6.995

Schneider, I. K., & Mattes, A. (2021). Mix is Different from Nix: Mouse Tracking Differentiates Ambivalence from Neutrality. *Journal of Experimental Social Psychology*, *95*, 104106. https://doi:10.1016/j.jesp.2021.104106

Schneider, I. K., van Harreveld, F., Rotteveel, M., Topolinski, S., van der Pligt, J., Schwarz, N., & Koole, S. L. (2015). The path of ambivalence: tracing the pull of opposing evaluations using mouse trajectories. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00996

Schneider, I. K., & Schwarz, N. (2017). Mixed feelings: the case of ambivalence. *Current Opinion in Behavioral Sciences*, *15*, 39–45. https://doi.org/10.1016/j.cobeha.2017.05.012

Schneider, I. K., Veenstra, L., van Harreveld, F., Schwarz, N., & Koole, S. L. (2016). Let’s not be indifferent about neutrality: Neutral ratings in the International Affective Picture System (IAPS) mask mixed affective responses. *Emotion*, *16*(4), 426–430. https://doi.org/10.1037/emo0000164

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2017). *Afex: Analysis of factorial experiments*. Retrieved from https://CRAN.R-project.org/packageafex

Singmann, H., & Kellen, D. (2013). MPTinR: Analysis of Multinomial Processing Tree models in R. *Behavior Research Methods*, *45*(2), 560—575. https://doi.org/10.3758/s13428012-0259-0

Stahl, C., & Aust, F. (2018). Evaluative Conditioning as Memory-Based Judgment. *Social Psychological Bulletin*, *13*(3). https://doi.org/10.5964/spb.v13i3.28589

Stahl, C., & Heycke, T. (2016). Evaluative Conditioning with Simultaneous and Sequential Pairings Under Incidental and Intentional Learning Conditions. *Social Cognition*, *34*(5), 382–412. https://doi.org/10.1521/soco.2016.34.5.382

Torchiano, M. (2016). *Effsize - a package for efficient effect size computation.* Retrieved from http://doi.org/10.5281/zenodo.1480624

van Dessel, P., de Houwer, J., Gast, A., & Tucker Smith, C. (2015). Instruction-Based Approach-Avoidance Effects. *Experimental Psychology*, *62*(3), 161–169. https://doi.org/10.1027/1618-3169/a000282 van Dessel, P., Mertens, G., Smith, C. T., & de Houwer, J. (2017). The Mere Exposure Instruction Effect. *Experimental Psychology*, *64*(5), 299–314. https://doi.org/10.1027/1618-3169/a000376

van Herreveld, F., Nohlen, H. U., & Schneider, I. K. (2015). The ABC of ambivalence: Affective, behavioral, and cognitive consequences of attitudinal conflict. *Advances in experimental social psychology*, *52*, 285-324.

Waroquier, L., Abadie, M., & Blaye, A. (2022). Age-related changes in controllable and uncontrollable processes underlying evaluative conditioning: A comparison between children ages 7 to 11 and young adults. *Developmental Psychology*, *58*(2), 359–366. https://doi.org/10.1037/dev0001296

Wickham, H. (2016). *ggplot2: elegant graphics for data analysis.* New York: Springer.

Zanon, R., De Houwer, J., Gast, A., & Smith, C. T. (2014). When does relational information influence evaluative conditioning? *Quarterly Journal of Experimental Psychology*, *67*(11), 2105–2122. https://doi.org/10.1080/17470218.2014.907324

Yonelinas, A. P., & Jacoby, L. L. (2012). The process-dissociation approach two decades later: Convergence, boundary conditions, and new directions. *Memory & Cognition*, *40*(5), 663–680. https://doi.org/10.3758/s13421-012-0205-5