Dissociating the roles of episodic retrieval and contingency awareness in valence contingency learning

Carina G. Giesen[1], Hannah Duderstadt[2], Jasmin Richter[3] & Klaus Rothermund[2]

[1]Health and Medical University Erfurt, Germany

[2]Friedrich Schiller University Jena, Germany

[3]University of Oslo, Norway

Cognition & Emotion, in press

ORCID

Carina G. Giesen 0000-0002-2395-4435

Jasmin Richter 0000-0002-3527-5923

Klaus Rothermund 0000-0002-9350-5272

Corresponding author: Carina G. Giesen, Health and Medical University Erfurt, Faculty of Health, Department of Psychology, Anger 64-73 99084 Erfurt, Germany. E-Mail: carina.giesen@hmu-erfurt.de. Hypotheses, design, and analyses were preregistered (https://osf.io/8u2bz). All materials, data, and analyses will be available at the Open Science Framework (https://osf.io/83vdr/.

Abstract

In the valence contingency learning task (VCT), participants evaluate target words which are preceded by nonwords. Nonwords are predictive for positive/negative evaluations. Previous studies demonstrated that this results in (a) reliable contingency learning effects, reflected in better performance for highly contingent nonword-valence pairings and (b) less reliable evaluative conditioning (EC) effects, reflected in more positive ratings of nonwords that were highly predictive of positive (vs. negative) evaluative responses. In a highly-powered (N=129) preregistered study, we investigated both effects and assessed whether they are a consequence of episodic retrieval of incidental stimulus-response (SR) episodes and/or propositional learning (indicated by contingency awareness). Participants were either explicitly instructed about contingencies (*instructed learning group*) or not (*incidental learning group*). Both groups then worked through the VCT, an explicit rating task, and a contingency awareness test. Both groups showed contingency learning effects and EC effects for nonwords. Multi-level analyses showed that controlling for previous SR co-occurrences fully accounted for contingency learning effects in the incidental learning group. In the instructed learning group, a residual effect of genuine valence contingency learning remained. Nonword-specific contingency awareness in turn fully accounted for EC effects in both learning groups, indicating that genuine contingency learning effects reflect propositional learning.

Keywords: contingency learning, evaluative conditioning, stimulus-response episodes, episodic retrieval, contingency awareness, propositional learning.

When we perceive a stimulus, we may respond in a multitude of ways. Often, however, our behavior is not as variable as we would assume. Instead, there is a high likelihood that our current response will equal the response we gave on the *last previous occurrence* of the very same stimulus. Put differently, if a stimulus repeats, it is likely that we will repeat the response we gave the last time when the stimulus was present. Why is this the case? In order to answer this question, one needs to know more about the cognitive processes that are at the heart of action control. These are captured by the *Binding and Retrieval in Action Control* (BRAC) framework (Frings et al., 2020). The BRAC framework stands in the tradition of *ideomotor theory* (e.g., Greenwald, 1970; Shin et al., 2010) and the influential *Theory of Event Coding* (TEC; Hommel et al., 2001). Accordingly, human action presumably relies on binding mechanisms that integrate mental representations of perceived stimuli (S) and executed responses (R) into SR episodes (*event files* or *bindings* are synonymous terms, see Frings, Beste et al., 2024). Event files are transient links between mental representations that are stored in episodic memory. They can be retrieved by repeating one of their elements. Retrieval of SR episodes often facilitates future actions by shortcutting action planning and selection stages.

To date, a burgeoning amount of evidence attests that storage and retrieval of transient episodic bindings are core mechanisms of action regulation. These mechanisms apply to a broad scope of stimuli, modalities, responses, and contexts (for an overview, see Frings et al., 2020; Henson et al., 2014; Kiesel et al., 2023). Although research on SR episodes is a prosperous endeavour, it is still not entirely clear whether and how binding and retrieval effects (which are mostly studied at the trial-to-trial level), relate to longer lasting learning effects (which are typically studied across blocks of trials). Although this is a much-debated topic (e.g., Frings, Foerster et al., 2024), systematic research findings are still scarce (see Moeller & Frings, 2017, for a detailed discussion). An idea that has received more and more attention in the recent past concerns the possibility that binding and learning effects are outcomes from related, possibly even identical processes (e.g., Schmidt et al., 2016, 2020; Giesen et al., 2020). In this view, one may regard both phenomena as representing two sides of the same coin. Transient SR episodes might represent a rudimentary short-term manifestation of learning (in the sense of one-trial learning). Similarly, learning effects could be regarded as the result of stimulus-based retrieval of recent SR episodes.

In the action control literature, existing evidence for the influence of episodic retrieval of recent SR episodes in producing learning effects mostly stems from the color-word contingency learning paradigm (Schmidt et al., 2007). In this task, participants classify the color of different words. Participants are not informed about the fact that the irrelevant word meaning is contingently mapped to color responses, as some words appear more often in a particular color than in all other colors. Learning of these word-response contingencies is indicated by faster performance on high contingent (frequent) compared with low contingent (infrequent) word-color key pairings. Giesen and colleagues (2020) created a variant of this paradigm. Participants classified the color of different words (adjectives) via key presses (red, green, blue, yellow). Each word was predictive for two color responses. Crucially, the paradigm allowed to analyze retrieval of SR episodes that emerged over consecutive trial sequences and contingency learning effects simultaneously, and robust evidence for both effects emerged in two experiments. More important, however, was the finding that in both experiments, contingency learning effects were effectively *eliminated* as soon as the authors added a predictor that accounted for effects that presumably stem from episodic retrieval of the most recent SR episode in a hierarchical multi-level regression approach. This supports the conclusion that color-word contingency learning effects can be attributed to episodic retrieval of recent SR episodes, triggered by the irrelevant word stimuli that are presented in a trial.

To bridge the gap between transient binding and retrieval effects and more permanent forms of learning, Giesen and colleagues (2020) proposed the *law of recency*, which posits that action control processes often rely on retrieval of previous SR episodes from memory. Specifically, episodic retrieval will reactivate the response that was executed during the *most recent episode* in which the stimulus was presented before. According to the *law of recency*, behaving in a particular way in a given situation increases the likelihood of demonstrating the same behavior *again* when the same situation is encountered the next time. One advantage of the *law of recency* is that it can be tested with the hierarchical multi-level modelling approach described above, which first tests for the presence of a learning effect. Thus, the contingency level of each trial is used as a single predictor for trial RT. Typically, RT will be faster for high contingency vs. low contingency trials. In a second step, *previous response* is added as a second predictor. This predictor codes which response was required the last time the identical nonword was presented. If the previous response matches the response that required on the current trial, performance is facilitated. If the previous response mismatches current trial requirements, performance is impeded. Thus, the *previous response* predictor accounts for systematic variance in the RT data that can be attributed to retrieval of transient SR episodes. The effect of the contingency factor after including the *previous response* factor as an additional predictor thus reflects genuine effects of contingency learning that predict responding over and above episodic retrieval. The critical question is then whether the learning effect remains significant or not. If the latter holds true, this implies that episodic retrieval of recent SR episodes fully accounts for the variance in the learning effect. If, on the other hand, a reliable residual effect of contingency learning remains even after introducing episodic retrieval of the last matching episode as a predictor into the regression, this would attest to the involvement of other processes reflecting abstract knowledge of the contingencies or other forms of cumulative learning in producing the contingency learning effect (Rudolph & Rothermund, 2024; Rudolph et al., in press; Xu & Mordkoff, 2020).

So far, the *law of recency* has been successfully applied to account for learning of stimulus-response contingencies (like word-color key responses, Giesen et al., 2020; Rudolph & Rothermund, 2024; Schmidt et al., 2020; Xu & Mordkoff, 2020), learning of response-response contingencies (Rothermund et al., 2024) as well as other phenomena that are otherwise attributed to processes of cognitive control (Güldenpenning et al., 2024; Rothermund et al., 2022). These studies are complemented by additional findings that highlight clear limits or boundary conditions under which the *law of recency* apparently does not apply. These examples stem from research on more complex forms of contingency learning effects. For instance, overshadowing-like effects (i.e., selective learning for more salient, compared to less salient cues) in contingency learning were mediated by contingency awareness (i.e., knowledge about the existing word-response contingencies underlying a particular task) rather than episodic retrieval of recent SR episodes (Arunkumar et al., 2022; see also Arunkumar, Rothermund & Giesen, 2024, Arunkumar, Rothermund, Kunde et al. 2024; Rudolph & Rothermund, 2024). It appears that once participants detect the underlying SR contingencies, applying this propositional knowledge can become another source that determines participants' behavior (De Houwer, 2014; Mitchell et al., 2009). Tentatively, this could imply that contingency learning (CL) effects can be the outcome of at least *two* different processes. One process is captured by episodic storage and retrieval of recent SR episodes which emerge by mere spatial-temporal contiguity. Although this mechanism of episodic retrieval contributes the largest share to the CL effect, it can be questioned whether this actually should be considered as learning proper. Episodic retrieval does not reflect any lasting change resulting from accumulated experiences with environmental regularities, which is a prerequisite for learning (De Houwer & Hughes, 2020). Instead, episodic retrieval according to the *law of recency* (Giesen et al., 2020) is limited to the influence of a single episode, and is undone and/or reversed by a new episode. The second process that may underlie CL effects is captured by the application of rule-based knowledge about predictive stimulus-stimulus or stimulus-response relationships. This knowledge can be due to inferences based on prior experience (Rudolph & Rothermund, 2024). However, this tentative conclusion is so far based on studies in which contingency awareness was only measured (except for a study by Rudolph & Rothermund, 2024, who provided participants with statements about color-word contingencies that were either correct or incorrect). To provide more substantial evidence for the interplay of contingency awareness, episodic retrieval processes, and learning phenomena, it would be advisable to experimentally manipulate contingency awareness.

Against this background, the theoretical motivation of the present study was twofold. First, we were interested whether we can generalize the findings from Giesen et al. (2020) to instances of evaluative learning (research aim 1). Second, we wanted to gain more evidence on the initial finding that awareness or knowledge about contingencies represents a boundary condition for the role of episodic retrieval of the last previous SR episode in producing CL effects (research aim 2). This insight would be informative, as the relation of binding and learning is still not well understood.

With respect to aim 1, the valence contingency task (VCT) is ideal (Schmidt & De Houwer, 2012a). In this task, participants evaluate target words which are preceded by nonwords. Two nonwords are mostly paired with (and therefore predictive of) targets that require a positive evaluative response, whereas two other nonwords are mostly paired with (and therefore predictive of) targets that require a negative evaluative response. According to Schmidt and De Houwer (2012a), this task produces robust CL effects, reflected in faster and more accurate performance for highly contingent nonword-valence pairings. Schmidt and De Houwer (2012a) also collected evaluative ratings for nonwords after the VCT, which reflected evaluative conditioning (EC) effects for nonwords, indicated by more positive ratings of nonwords that were highly predictive for positive (vs. negative) evaluative responses. Contingency awareness mediated both and produced stronger CL effects in the VCT (errors but not RT) and EC effects (note that Gast et al., 2020, obtained the same pattern of results in two highly powered replication studies only for the VCT, but found no EC effects; we come back to this issue in the General Discussion).

With respect to aim 2, one first needs to address to what extent contingency awareness impacts on contingency learning. For reasons of consistency and brevity, we limit our focus only on color-word contingency learning. Indeed, some studies addressed the role of contingency awareness on color-word contingency learning. According to Schmidt et al. (2007), color-word contingency learning effects typically emerge irrespective of whether participants are explicitly aware of experienced contingencies or not (Schmidt et al., 2007). That is, learning effects emerge for participants who can later report the SR contingencies (aware participants), but also for those who perform at chance level when asked to report SR contingencies (unaware participants). Nevertheless, once participants become aware of color-word contingencies, this is beneficial, meaning that contingency learning effects are stronger for aware participants (Schmidt & De Houwer, 2012b; Rudolph & Rothermund, 2024). This pattern holds for studies that only measure contingency awareness at the end of the task (e.g., Schmidt et al., 2007), but also for studies that endorse experimental manipulations of contingency awareness (Schmidt & De Houwer, 2012b).

For instance, Schmidt and De Houwer (2012b, Experiment 1) informed one group of participants about the actual contingencies underlying the valence contingency learning task (instructed learning group). A second group of participants did not receive this information (incidental learning group). CL effects were stronger for the instructed than for incidental learning group; moreover, these effects were mediated by subjective contingency awareness. In light of the previously discussed finding that (a) adding a previous response predictor to capture systematic effects of retrieval of SR episodes completely accounted for CL effects in this task (Giesen et al., 2020) and the finding that (b) contingency awareness eliminated the role of SR retrieval in other learning effects (Arunkumar et al., 2022), the question arises whether CL effects can sometimes be based on retrieval of SR episodes and sometimes on propositional learning that is reflected in contingency awareness.

In our view, the VCT is well suited to dissociate the roles of episodic retrieval of recent SR episodes and contingency awareness in producing learning effects. To fulfill our proposed research aims, we used Experiment 2 by Schmidt and De Houwer (2012a). However, to investigate whether contingency awareness represents a boundary condition for episodic retrieval processes, we added an independent factor to the design by experimentally manipulating contingency awareness, similar to Schmidt and De Houwer (2012b, Experiment 1). Having two learning groups is vital, because this approach represents an experimental manipulation of contingency awareness that is superior to other approaches of studying the influence of contingency awareness on learning, as for instance measures of contingency awareness that face problems of their own (see Discussion section). Participants were either explicitly instructed about contingencies (*instructed learning group*) or not (*incidental learning group*). Both groups then worked through the VCT, an explicit rating task, and a contingency awareness test. In line with Schmidt and De Houwer (2012a), we expected effects of contingency learning in the VCT (reflected in better performance on high vs. low contingent nonword-valence pairings) as well as more positive evaluations of nonwords that were predictive of positive (vs. negative) valence responses (EC effects). Furthermore, we then planned to run multi-level analyses on contingency learning effects in the VCT to investigate whether contingency learning effects can be explained by episodic retrieval of recent SR episodes. Based on previous findings (Arunkumar et al., 2022, Arunkumar, Rothermund & Giesen, 2024, Arunkumar, Rothermund, Kunde et al. 2024; Giesen et al., 2020), we expected that adding a *previous response* predictor to account for episodic retrieval effects would eliminate contingency learning effects, but only in the *incidental learning group*. In the *instructed learning group*, we expected that contingency learning effects would remain significant, even when the *previous response* predictor was added. Hypotheses, design, and analyses were preregistered (https://osf.io/8u2bz). All materials, data, and analyses will be available (pending acceptance for publication) at the Open Science Framework (https://osf.io/83vdr/; view-only link for peer review: https://osf.io/ycwxv/?view_only=e73929a5a37b4a0486e3326d140cfe74).

Method

Required sample size, preregistration, and ethical approval

According to Brysbaert (2019), a sample size of N=130 participants is required to detect an effect size of d=0.5 with a statistical power of 1-ß=.80 in a two-tailed independent samples t-test with α=0.05 to detect a significant difference in contingency learning effects between groups. In accordance with guidelines of the American Psychological Association, no ethics approval was required because no cover-story or misleading or suggestive information was conveyed to participants. This procedure is in accordance with the ethical standards at the Institute of Psychology at the University Jena, where the study was planned, organized, and conducted.

Participants

In total, 130 native German speaking participants were recruited online via Prolific Academic, using a desktop computer or laptop. The study duration was 25 minutes. For one participant, the program crashed during data collection and no data were saved. Hence, data of 129 participants were analyzed (M_{age}= 32.6, range: 18-70 years, 62 female, 65 males, 2 diverse). Participants received £3.75 as compensation. Explicit informed consent to participate was collected electronically at the beginning of the study.

Material

The study was programmed in E-Prime 3.0 and converted for online data collection with E-Prime Go. All stimuli from Schmidt and De Houwer (2012a, Experiment 2) were used, including four nonwords (i.e., words not existing in German: *nijaron, fevkani, kadirga, lokanta*) and 24 valenced target words, translated to German. Twelve of these were positive and 12 were negative (see Table 1). All words were presented in white font on a black screen.

Design

The study comprised a 2×2×2 mixed factors design with the between-subject factor group (*instructed learning*, *n*=65 vs. *incidental learning*, *n*=64) and the within-subject factors contingency (high vs. low), and valence (positive vs. negative). Contingency was manipulated by presenting two nonwords eight times with positive target words and two times with negative target words; the other two nonwords were presented eight times with negative and two times with positive target words. Hence, we used a 4:1 contingency ratio. Valence indicated whether nonwords were predictive for positive vs. negative target words; assignment of nonwords to each valence level was counterbalanced across participants.

Procedure

After providing informed consent, participants were randomly assigned to one of the group conditions (*instructed* vs. *incidental learning*). All participants were informed that they would first see a nonword on screen, followed by a valent target word, which would appear either above or below the nonword. Participants were instructed to classify the valence of the target word by pressing a left (F) or right (J) key as fast and as correctly as possible. Assignment of valence to response keys was counterbalanced between participants. Participants in the *incidental learning group* then continued with the practice block (Figure 1 provides an overview of the sequence of procedural events during the experiment for each group). Participants in the *instructed learning* group were explicitly informed that two nonwords would occur most often with positive targets, whereas the other two nonwords would occur most often with negative target words. We did not reveal which exact nonword would be predictive for each valence, though.

Then, the valence contingency learning task started with a brief practice block (40 trials sampled randomly from the main experimental block). Participants continued to the main block only if they produced less than 50% errors during practice; otherwise, the experiment was ended. All participants passed the accuracy criterion.

After practice, only participants in the *instructed learning* group were prompted with a first objective contingency awareness test. In past studies in our lab, we noticed that probing participants' awareness prior to a task renders them more likely to actively attend to contingencies. Every nonword was presented in isolation. Participants were told to report whether the nonword was predictive of positive or negative target words in the preceding practice block by pressing P or N, respectively, and further reported how confident they were on a scale from 1 (very uncertain) to 5 (very certain). Participants in the *incidental learning group* received no contingency awareness test and directly proceeded from practice to the main experiment.

Identical to Schmidt and De Houwer (2012a), the main block comprised 480 trials that were constructed with respect to the factorial design. Every nonword was presented 120 times. Two nonwords were presented eight times with each of the 12 positive target words and two times with each of the 12 negative target words; the other two nonwords were presented eight times with every negative target word and two times with every positive target word. Hence, 384 trials (80%) represented high contingency nonword-valence pairings, and 96 trials (20%) represented low contingency nonword-valence pairings. Trial presentation was randomized. After 240 trials, participants were given a brief self-paced break (until response to continue).

The trial procedure was also identical to Schmidt and De Houwer (2012a): Every trial during the practice and main block started with a centrally presented fixation cross (250 ms), followed by a blank screen (50 ms). Then, a nonword appeared either above or below the screen center (250 ms; nonword position was balanced within participants). Then, a target word appeared at the other position either below or above the screen center (until response or until a maximal duration of 2000 ms was reached, which would end the screen and count as nonresponse). Participants then had to classify the valence of target words as fast as possible by pressing J or F. Correct responses were followed by another blank screen for 500 ms; erroneous or too slow responses were followed by the feedback message “Fehler, falsche Reaktion” (“error, wrong response”) displayed in red font for 1000 ms. Then, the next trial started.

After the main block, a rating task followed, like in Schmidt and De Houwer (2012a). In this task, all four nonwords as well as two positive (flower, hug) and two negative (guns, crime) words were presented in randomized order. Participants had to rate the valence of each word on a scale from 1 (negative) to 8 (positive). Ratings for nonwords allowed to assess EC effects for nonwords. Next, all participants were informed that two nonwords were predictive for positive targets, whereas two other nonwords were predictive for negative targets. To assess subjective awareness of contingencies, participants were asked whether they noticed this relationship (yes/no responses were collected). Then, another objective contingency awareness test followed. Objective contingency awareness was assessed for all nonwords. Participants were told to report whether they felt a given nonword was mostly paired with positive or negative target words by pressing P or N, followed by a confidence rating on a scale from 1 (very uncertain) to 5 (very certain). At the end of the study, all participants were debriefed and compensated accordingly.

Data preparation

In accordance with the preregistration, for the valence contingency learning task, erroneous trials or trials that were preceded by an erroneous trial were removed (9.7%) as well as trials with outlier RT[1] (5.1%). For multilevel analyses, trial RT were normalized using inverse normal transformation.

Results

For reasons of brevity, standard analyses of CL effects (i.e., without partialling out previous response effects) and their correlations with contingency awareness are reported in the Appendix. In sum, we obtained (a) CL effects in the VCT (reflected in faster and more accurate performance for high vs. low contingent nonword-valence pairings) as well as (b) EC effects for nonwords (reflected in more positive evaluations for nonwords that were predictive of positive compared with negative valence). Importantly, both CL and EC effects were significant already in the incidental learning group and therefore directly replicated Schmidt and De Houwer (2012a). Notably, both CL and EC effects were more pronounced in the instructed learning group, which is

---

[1] Reaction times below 150 ms or more than 1.5 interquartile ranges above the 75th percentile of the individual RT distribution were regarded as outliers (Tukey, 1977).

in line with findings by Schmidt and De Houwer (2012b). In the following, we will present additional analyses (preregistered and exploratory) that go beyond the analyses of Schmidt and De Houwer (2012a) and that aimed to dissociate the roles of previous response retrieval and contingency awareness in CL effects.

Preregistered analyses

Controlling for the last previous stimulus-response occurrence in CL effects. We investigated whether previous stimulus-response co-occurrences affected task performance and whether this accounted for systematic variance in valence contingency learning effects (cf. Giesen et al., 2020; Güldenpenning et al., 2024; Rudolph & Rothermund, 2024). RT data were analyzed with a linear mixed-effects model (LMM), using trials as units of analysis (level-1 predictor), nested within participants (level-2 predictor to account for dependencies between trial performance). We computed four random intercept models in hierarchical order with inverse normal-transformed trial RT as dependent measure and participants as random effects. Predictors for each model were centered via contrast coding of each predictor with respect to the relative frequencies of its factor levels, see descriptions below for details. To account for systematic variance due to stimulus-based episodic retrieval (Model 2), every trial was referenced back to the last previous occurrence of the respective nonword. Then, the relation between the response required in the current trial and the last previous occurrence was coded (previous response same vs. different as current response). Thus, only trials with stimulus repetition at the level of nonwords were considered in all Models (see Table 2 for an overview of results).

In Model 1, Contingency (contrast coded[2] with *high contingency* = -.20 and *low contingency* =.80) was the only predictor, which produced a significant CL effect, b=0.03, t(52637)=9.619, p<.001. For Model 2, the predictor *previous response* was added. This predictor coded the relation between responses of the current trial and the last previous occurrence of a correct response, which could be the same (i.e., previous response = same as current response, contrast weight: -0.3195; or previous response = different, contrast weight: 0.6715). Adding *previous response* as an additional predictor in Model 2 produced a significant effect for this factor, b=0.041, t(52638)=12.46, p<.001, indicating faster performance when the current response matched (vs. mismatched) the previous response. Importantly, including *previous response* as an additional predictor rendered the effect for the contingency predictor nonsignificant, b=0.007, t(52637)=1.779, p=.075. For Model 3, we coded how distant the last previous occurrence of each nonword was from the current trial (lag-1 to lag-30); distance represented a continuous predictor and was log-transformed and centered within participants for analyses. Model 3 yielded the same effect pattern as Model 2 as well as a main effect of distance, b=0.018, t(52637)=11.47, p<.001, qualified by an interaction of distance × previous response, b=-0.053, t(52638)=-15.42, p<.001. Follow-up tests showed that the benefit of same vs. different previous response trials was larger for immediate repetitions from the previous trial (i.e., distance = lag-1), Δ=38 ms, t(128)=13.23,

[2] All predictors indicating a contrast between two conditions were coded to have (1) a mean of zero across all trials within the analysis, and (2) a difference of 1 between the two weights. The general formulas that will satisfy these standards are: $w_1 = \frac{f_2}{(f_1+f_2)}$ and (2) $w_2 = \frac{-f_1}{(f_1+f_2)}$, where w_1 and w_2 are the regression weights that define a contrast, and f_1 and f_2 are the number of trials per condition (see Rudolph & Rothermund, 2024). Equal trial frequencies in each condition will result in a contrast coding of w_1 = .5 and w_2 = -.5. If the frequencies are unequal, however, this will result in unequal weights, so that the condition with the higher frequency will be assigned with a lower (absolute) weight (and vice versa). Thus, the resulting regression coefficient reflects the difference between the two conditions (in milliseconds), and the main effects and interactions of the predictors can be interpreted simultaneously.

p<.001, d_z=1.16, than for non-immediate repetitions (distance of lag-2 to lag-30), Δ=4 ms, *t*(128)=3.08, *p*=.003, d_z=0.27. In the final Model 4, group was added as an additional predictor (instructed learning = 0.5, incidental learning = -0.5). Model 4 yielded the same effect pattern as Model 3 but also revealed an interaction of contingency and group, *b*=0.025, *t*(52638)=2.56, *p*=.010, indicating that CL effects were larger in the instructed vs. incidental group; no other effect of group was significant. To follow up on this interaction, we repeated the LMM analyses (Model 1 to Model 3) separately for the incidental and instructed learning groups.

For the *incidental learning* group, results were comparable to the overall analyses. In detail, Model 1 revealed a significant effect of contingency, *b*=0.021, *t(*25985*)*=4.596, *p*<.001. In Model 2, the contingency predictor was no longer significant, *b*=-0.004, *t(*25985*)*=0.686, *p*=.493, as soon as previous response was considered, *b*=0.041, *t(*25985*)*=8.934, *p*<.001. Model 3 produced a similar result pattern plus a main effect of distance, *b*=0.016, *t(*25985*)*=7.452, *p*<.001, as well as the distance × previous response interaction, *b*=-0.06, *t(*25985*)*=-12.696, *p*<.001.

For the *instructed learning* group, a different result pattern emerged. In detail, Model 1 revealed a significant effect of contingency, *b*=0.04, *t*(26652)=8.86, *p*<.001. In Model 2, this effect *remained* significant, *b*=0.017, *t*(26652)=3.065, *p*=.002 (albeit somewhat reduced in size), even when previous response was added, which was also a significant predictor, *b*=0.042, *t*(26652)=8.702, *p*<.001. Model 3 yielded significant effects of contingency, *b*=0.017, *t*(26652)=3.112, *p*=.001, previous response, *b*=0.041, *t*(26652)=8.490, *p*<.001, an additional effect of distance, *b*=0.02, *t*(26652)=8.720, *p*<.001, and the distance × previous response interaction, *b*=-0.04, *t*(26652)=-9.285, *p*<.001.

Hence, whereas previous response effects completely statistically accounted for the CL effects for the *incidental learning group* and rendered them nonsignificant, this was not the case for the *instructed learning group*. Here, CL effects remained significant even when the *previous response* predictor was added. This pattern suggests that valence contingency learning is – at least in parts – the outcome of other processes and cannot be solely accounted for by episodic retrieval of previous SR episodes. Given that (a) participants in the *instructed learning group* were also characterized by a higher level of contingency awareness and since (b) existing studies already demonstrated that contingency awareness is a strong moderating influence (Schmidt & De Houwer, 2012b) that also limits the explanatory power of episodic retrieval processes in accounting for contingency learning (Arunkumar et al., 2022, 2024a, 2024b), we reasoned that learning effects in this group are due to propositional learning as reflected in contingency awareness. We therefore conducted additional analyses to explore the role of contingency awareness for the emergence of evaluative learning effects (cf. Rudolph & Rothermund, 2024).

Exploratory analyses

Predictors of evaluative conditioning effects. In the present study, participants showed EC effects after working through the VCT. Given that EC effects emerged in both groups, we were interested in whether EC effects could be statistically accounted for by stimulus-specific contingency awareness, that is, explicit memory for high contingency nonword-valence pairings (Baeyens et al., 1990; Pleyers et al., 2007; Stahl & Unkelbach, 2009). We ran another LMM analysis on evaluative ratings for each stimulus (level-1 predictor), nested within participants (level-2 predictor). All models represented random intercept models. Only significant effects are discussed. In Model 1, valence of the high contingency nonword-target pairings (positive valence: +0.5 vs. negative valence: -0.5) was the only predictor and produced a significant effect, b=.96, $t(516)$=6.841, p<.001; the EC effect. In Model 2, we added two additional predictors, namely group (instructed learning=+0.5, incidental learning=-0.5) as well as stimulus-specific contingency awareness (correct reports of the high contingency nonword-valence pairing for a specific nonword were coded as 0.3275, inaccurate reports were coded as -0.6724). This final model yielded a significant main effect of valence, b=.91, t(588.5)=6.928, p<.001, which was qualified by several interactions, namely between valence × contingency awareness, b=1.95, t(464.5)=6.895, p<.001, and between valence × group, b=0.65, t(588.5)=2.459, p=.014, which were further qualified by a three-way interaction between valence × contingency awareness × group, b=1.41, t(464.5)=2.491, p=.013, meaning that (a) EC effects were limited to those stimuli for which high contingency valence pairings were accurately reported (blue lines in Figure 2), but were absent otherwise; furthermore, (b) this effect was more pronounced for participants in the instructed learning group (compare slope of blue lines for left and right side of Figure 2). No other effects were significant.

Discussion

In the present study, we employed the valence contingency learning task, developed by Schmidt and De Houwer (2012a). Participants classified the valence of different target words via keypress. Target words were preceded by nonwords. Importantly, two nonwords were predictive of positive evaluative responses to targets, whereas two other nonwords were predictive of negative evaluative responses to targets. Participants were either instructed about these nonword-valence contingencies (instructed learning group) or not (incidental learning group). Results were in line with our preregistered expectations. In detail, we obtained (a) CL effects in the VCT (reflected in faster and more accurate performance for high vs. low contingent nonword-valence pairings) as well as (b) EC effects for nonwords (reflected in more positive evaluations for nonwords that were predictive of positive compared with negative valence). Importantly, both CL and EC effects were significant already in the incidental learning group and therefore directly replicated Schmidt and De Houwer (2012a). Notably, both CL and EC effects were more pronounced in the instructed learning group, which is in line with findings by Schmidt and De Houwer (2012b).

To investigate our central question on how retrieval of the last previous stimulus-response occurrence and propositional learning (manipulated via learning instructions) affect CL effects, we then ran several multi-level analyses. Put differently, this analysis explores whether CL effects can be statistically accounted for by stimulus-driven episodic retrieval processes of recent SR episodes. In line with our preregistered expectations, CL effects were eliminated by adding the *previous response* predictor – but only in the *incidental learning group*. For the *instructed learning group*, CL effects remained significant, even after the *previous response* predictor was added (cf. Rudolph & Rothermund, 2024). This suggests that valence contingency learning is – at least in parts – the outcome of propositional learning (manipulated via learning instructions) and cannot be solely accounted for by episodic retrieval of previous SR episodes. For participants of the incidental learning group, it appears that contingency learning effects are due to episodic retrieval of the response that was executed the last time the same stimulus was presented. For participants of the instructed learning group, contingency learning effects can be the result of two processes, namely episodic retrieval and/or the strategic application of stimulus-specific contingency rules that were detected by the participants. Note that we did not instruct participants in this group about the presence of specific nonword-valence contingencies but only about the general presence of nonword-valence contingencies. Under such conditions of instructed learning, explicit knowledge for nonword-valence contingencies is more likely to arise and thus may also systematically affect performance.

Theoretical implications

Our findings bear several important theoretical implications. First, finding CL and EC effects in the incidental learning group can be understood as a direct replication of Experiment 2 by Schmidt and De Houwer (2012a), albeit in an online study with German participants. Our findings therefore illustrate that the effects generalize across different sample characteristics (e.g., stimulus material, nationality, data collected in lab vs. online). This insight is important, as Gast and colleagues (2020) could only partially replicate findings from the very same experiment in a direct replication attempt with multiple experiments. Their study successfully replicated valence contingency learning effects for nonwords but failed to replicate effects of contingencies on evaluative ratings (EC). According to Gast and colleagues, this could be the result of a coding error that possibly resulted in an overestimation of awareness-independent processes in the data by Schmidt and De Houwer (2012a; see also the Erratum). This argument is important as contingency awareness (in the sense of explicit memory for frequent/high contingency nonword-valence pairings) is discussed as a necessary precondition for the emergence of EC effects (Gast, 2018; Hofmann et al., 2010).

Second, our findings can help to reconcile the findings by Schmidt and De Houwer (2012a) and Gast and colleagues (2020) and therefore fill an important gap in the existing literature. For this, the exploratory analysis on evaluative ratings is particularly informative: The analysis indicated that EC effects in both groups were a function of nonword-specific (i.e., CS-specific) contingency awareness, as EC effects emerged only for those nonwords for which a given participant could correctly report the high contingency valence category. In turn, for nonwords for which the highly contingent valence category could not be correctly reported, EC effects were absent. The results of this exploratory multi-level analysis converge with findings by Gast and colleagues (2020) who also reported that EC effects are dependent on item-specific memory for nonword-valence pairings. Note that Schmidt and De Houwer (2012a, Erratum) also reported a positive correlation between objective contingency awareness and EC effects. Together, these findings suggest two conclusions: (1) Whereas contingency learning effects within the VCT may in part occur independently of contingency awareness via retrieval of SR episodes, (2) to manifest in other evaluative learning effects like a change of liking of nonwords, item-specific awareness of contingent valence pairings is a necessary precondition (De Houwer, 2018; Förderer & Unkelbach, 2013; Lovibond & Shanks, 2002; Mitchell et al., 2009; Stahl et al., 2009).

Third, the fact that previous responses statistically fully accounted for variance in contingency learning in the VCT for the incidental learning group indicates that the *law of recency* not only holds for color-word contingency learning, but also generalizes to other contingency learning paradigms. Hence, this result can be taken as another piece of evidence in favour of understanding contingency learning as resulting from the retrieval of discrete SR episodes (Giesen et al., 2020) rather than from cumulative learning effects represented in abstract form – at least when particular conditions are met.

Novelty and Impact of the Present Findings

To our knowledge, this is the first study that has systematically investigated whether effects of valence contingency learning can be attributed to episodic retrieval of transient SR episodes. Moreover, this is the first study to investigate effects of genuine valence contingency learning (i.e., after accounting for the effects of episodic response retrieval). To demonstrate this, our study used a new analytical approach that goes beyond the analyses provided by Schmidt and De Houwer (2012a). Thus, our approach is novel and important, as (valence) contingency learning was speculated to reflect an instance of implicit learning – that is, learning in the absence of awareness. This reasoning was based on data that indicated that contingency learning effects emerge regardless of whether participants can later report the underlying contingencies or not (Schmidt et al., 2007; Schmidt & De Houwer, 2012a). Especially the latter case was taken as evidence for implicit learning. The findings by Giesen et al. (2020) as well as the present findings illustrate that this understanding is incorrect, as a large part of the performance pattern that was initially interpreted as a genuine effect of contingency learning actually reflects episodic retrieval of the last response – that is, a transient episodic linkage between stimuli and responses that bears *no underlying contingency*. Along with De Houwer and Hughes (2020), we define learning as “an observable change in the behaviour of a specific organism as a consequence of regularities in the environment of that organism” (p. 4). Transient SR episodes, however, do not conform to this definition, as they do not represent contingent SR relationships, but merely arbitrary and transient linkages. Unlike the study by Giesen et al. (2020), the present study demonstrates this for valence contingency learning in a design in which awareness was not only measured, but also experimentally manipulated.

Importantly, our study also demonstrates genuine valence contingency learning effects after controlling for (i.e., partialling out) episodic retrieval effects. The remaining residual valence contingency learning effect is shown to depend on contingency awareness because it was larger and only present when participants were instructed to look for stimulus contingencies in the task. None of these findings has been reported in a previous study. Our findings are therefore not consistent with the view that binding and learning effects result from the same underlying mechanism. Instead, our findings suggest that episodic retrieval and propositional knowledge shape behaviour independently and via separate routes. This is an important and novel contribution to the question regarding the relation between SR binding and retrieval and learning. Future research is needed to more closely investigate the relationship between both processes.

Limitations

The present study used an experimental manipulation of contingency awareness that should affect learning of contingencies differently for each group. However, contingency awareness was only assessed at the end of the study. Hence, we do not know whether our measure of contingency awareness adequately captures the amount of contingency awareness in participants *during* the task (i.e., the "immediacy criterion" according to Lovibond & Shanks, 2002). For instance, it is possible that some participants were aware of the contingencies during the task but then forgot them partially or completely over the remaining course of the study. Alternatively, however, participants might have become aware of the contingencies only during later phases of the experiment, so that contingency awareness assessed at the end of the experiment might not have been present during the entire study. We want to point out that the present measures of episodic retrieval and contingency awareness were taken at different time points (the former took place on-task, the latter at the end of the study). Hence, our measure of contingency awareness might either under- or overestimate the actual amount of contingency awareness that participants developed during the task. Relatedly, our findings do not preclude the possibility that previous response effects depended on some kind of awareness of SR links during the learning task. Specifically, previous response effects may depend on participants' explicitly remembering which response they gave on the last occurrence of the same stimulus (but see Giesen & Rothermund, 2015; Rudolph et al., 2024). Knowing the previous response to a given stimulus during the VCT task may however only partially overlap with correctly reporting the highly contingent valence for a nonword after the learning phase. Relatedly, our conclusion that contingency awareness influences the strength of (valence) CL effects does not imply that contingency awareness is a necessary condition for the emergence of a genuine CL effect. The latter conclusion depends on the interpretation of null effects, and should be taken with caution. Although we did not find a significant residual CL effect in the incidental learning group, for which (measured) contingency awareness was low, this does not rule out the possibility that such an effect exists and could be obtained under different conditions (e.g., with a much larger sample, or with a stronger contingency manipulation). Considering these issues, it would be premature to interpret CL effects in the absence of post-learning contingency awareness as evidence for unaware CL learning. To address this issue, one would have needed an *on-task* measure of contingency awareness that takes place during the VCL task (cf. Giesen & Rothermund, 2015; Lovibond & Shanks, 2002). However, one potential pitfall of such an improved contingency measure is that it cannot be applied unobtrusively in all learning conditions, as occasional on-task assessment of contingency awareness could render participants in the incidental learning group suspicious. That is, it would effectively counteract the incidental learning condition and induce participants to search for contingencies.

Throughout the study, we interpreted effects of previous responses as an indicator for episodic retrieval of transient SR episodes/event files as stated in the *law of recency* (Giesen et all., 2020). However, we want to point out that this is by no means the first or only account that considers the power of episodic retrieval processes in producing learning effects. Indeed, there is a rich body of instance-based learning accounts or exemplar memory models (as for instance the Parallel Episodic Processing [PEP] model, Schmidt et al., 2016; the instance theory of automatization, Logan, 1988; or the MINERVA model, e.g., Mitchell et al., 2009). We want to point out that the present study is not well suited to dissociate among different exemplar models, but this was never the aim of our research. The benefit from the *law of recency* or the BRAC framework is based on the fact that they bridge the gap between research fields that became more and more independent and consequently less interconnected. By grounding findings from learning literature in the terminology of action control research, the BRAC framework sets the focus on possible common underlying mechanisms (like storage and retrieval of transient SR episodes), rather than explaining effect- or paradigm-specific idiosyncrasies. However, our findings may also be consistent with other processes that do not invoke the concept of episodic retrieval. Future research is needed to explore this question in more detail.

Conclusion

Our study showed that effects of nonword-valence contingencies on evaluation times are mediated to a large extent by episodic retrieval processes. On top of that, propositional learning elicited by instructed contingencies had a causal influence on the resulting CL effect, and also explained effects of nonword-valence contingencies on evaluative ratings (EC effect). These findings support an account of CL effects in terms of two underlying sources: Episodic retrieval, which should not be considered as learning proper, and propositional knowledge as reflected in subjective awareness of environmental contingencies.

Acknowledgements

We thank Julius Protte for his help in programming the experiments. Data were collected as part of Hannah Duderstadt's Master thesis. The study was funded by grants from the German Research Foundation to Carina G. Giesen (GI 1295/2-1,2) and Klaus Rothermund (RO 1272/13-1,2).

Declaration of Interests

The authors report there are no competing interests to declare.

References

Arunkumar, M., Rothermund, K., Kunde, W., & Giesen, C. G. (2022). Being in the know: The role of awareness and retrieval of transient stimulus-response bindings in selective contingency learning. *Journal of Cognition*, *5*(1), 36. https://doi.org/10.5334/joc.227

Arunkumar, M., Rothermund, K., & Giesen, C. G. (2024). One link to link them all: Indirect response activation through stimulus-stimulus associations in contingency learning. *Experimental Psychology*, *70*(5), 259–270. https://doi.org/10.1027/1618-3169/a000597

Arunkumar, M., Rothermund, K., Kunde, W., Mocke, V., & Giesen, C. G. (2024). Cling together, swing together? Assessing indirect retrieval of stimulus-response bindings for associated stimuli. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-024-02525-0

Aust, F., Haaf, J. M., & Stahl, C. (2019). A memory-based judgment account of expectancy-liking dissociations in evaluative conditioning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(3), 417–439. https://doi.org/10.1037/xlm0000600

Baeyens, F., Eelen, P., & van den Bergh, O. (1990). Contingency awareness in evaluative conditioning: A case for unaware affective-evaluative learning. *Cognition & Emotion*, *4*(1), 3–18. https://doi.org/10.1080/02699939008406760

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition.*, *2*(1), 1–38. https://doi.org/10.5334/joc.72

De Houwer, J. (2014). Why a propositional single-process model of associative learning deserves to be defended. In J. W. Sherman & B. Gawronski (Eds.), *Dual processes in social psychology* (pp. 530-541). New York: Guilford.

De Houwer, J. (2018). Propositional models of evaluative conditioning. *Social Psychological Bulletin*, *13*(3). doi:10.5964/spb.v13i3.28046

De Houwer, J., & Hughes, S. (2020). The psychology of learning: An introduction from a functional-cognitive perspective. Cambridge, MA: The MIT Press.

Förderer, S., & Unkelbach, C. (2013). On the Stability of Evaluative Conditioning Effects. *Social Psychology*, *44*(6), 380–389. https://doi.org/10.1027/1864-9335/a000150

Frings, C., Beste, C., Benini, E., Möller, M., Dignath, D., Giesen, C. G., Hommel, B., Kiesel, A., Koch, I., & Kunde, W. (2024). Consensus definitions of perception-actionintegration in action control. *Communications Psychology*, *2*(1). https://doi.org/10.1038/s44271-023-00050-9

Frings, C., Foerster, A., Moeller, B., Pastötter, B., & Pfister, R. (2024). The relation between learning and stimulus–response binding.*Psychological Review, 131*(5), 1290–1296. https://doi.org/10.1037/rev0000449

Frings, C., Hommel, B., Koch, I., Rothermund, K., Dignath, D., Giesen, C., Kiesel, A., Kunde, W., Mayr, S., & Moeller, B. (2020). Binding and retrieval in action control (BRAC). *Trends in Cognitive Sciences*, *24*(5), 375–387. https://doi.org/10.1016/j.tics.2020.02.004

Gast, A., Richter, J., & Ruszpel, B. (2020). Is there evidence for unaware evaluative conditioning in a valence contingency learning task? *Cognition and Emotion*, *34*(1), 57–73. https://doi.org/10.1080/02699931.2019.1652147

Giesen, C., & Rothermund, K. (2015). Adapting to stimulus–response contingencies without noticing them. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(6), 1475–1481. https://doi.org/10.1037/xhp0000122

Giesen, C. G., Schmidt, J. R., & Rothermund, K. (2020). The law of recency: An episodic stimulus-response retrieval account of habit acquisition. *Frontiers in Psychology*(10), Article 2927, 2927. https://doi.org/10.3389/fpsyg.2019.02927

Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideo-motor mechanism. *Psychological Review*, *77*(2), 73–99. https://doi.org/10.1037/h0028689

Güldenpenning, I., Böer, N. T., Kunde, W., Giesen, C. G., Rothermund, K., & Weigelt, M. (2024). Context-specific adaptation for head fakes in basketball: A study on player-specific fake-frequency schedules. *Psychological Research.* Advance online publication. https://doi.org/10.1007/s00426-024-01977-2

Henson, R. N., Eckstein, D., Waszak, F., Frings, C., & Horner, A. J. (2014). Stimulus–response bindings in priming. *Trends in Cognitive Sciences*, *18*(7), 376–384. https://doi.org/10.1016/j.tics.2014.03.004

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

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*(5), 849–937. https://doi.org/10.1017/S0140525X01000103

Kiesel, A., Fournier, L. R., Giesen, C. G., Mayr, S., & Frings, C. (2023). Core mechanisms in action control: Binding and retrieval. *Journal of Cognition*, *6*(1), 2. https://doi.org/10.5334/joc.253

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*(4), 492–527. https://doi.org/10.1037/0033-295X.95.4.492

Lovibond, P. F., & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behavior Processes*, *28*(1), 3–26. https://doi.org/10.1037/0097-7403.28.1.3

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*(2), 183–198. https://doi.org/10.1017/S0140525X09000855

Moeller, B., & Frings, C. (2017). Dissociation of binding and learning processes. *Attention, Perception, & Psychophysics*, *79*(8), 2590–2605. https://doi.org/10.3758/s13414-017-1393-7

Pleyers, G., Corneille, O., Luminet, O., & Yzerbyt, V. (2007). Aware and (dis)liking: Item-based analyses reveal that valence acquisition via evaluative conditioning emerges only when there is contingency awareness. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(1), 130–144. https://doi.org/10.1037/0278-7393.33.1.130

Rothermund, K., Gollnick, N., & Giesen, C. G. (2022). Accounting for proportion congruency effects in the Stroop task in a confounded setup: Retrieval of stimulus-response episodes explains it all. *Journal of Cognition*, *5*(1), 39. https://doi.org/10.5334/joc.232

Rothermund, K., Sprengholz, P., & Moeller, B. (2023). *Contingency learning of response sequences: Disentangling effects of episodic retrieval, association formation, and propositional knowledge*. MS in preparation.

Rudolph, M., Giesen, C. G., & Rothermund, K. (in press). False contingency knowledge reverses the color-word contingency learning effect. *Journal of Experimental Psychology. Learning, Memory, and Cognition.*

Rudolph, M., Giesen, C. G., & Rothermund, K. (2024). *Stimulus-response binding and retrieval operates independently of contingency awareness. A mega-analysis.* MS submitted for publication.

Rudolph, M., & Rothermund, K. (2024). Two sources of color-word contingency learning: Episodic retrieval of stimulus-response bindings and propositional knowledge. *Journal of Experimental Psychology. Learning, Memory, and Cognition.* Advance online publication. https://doi.org/10.1037/xlm0001353

Schmidt, J. R., Crump, M. J. C., Cheesman, J., & Besner, D. (2007). Contingency learning without awareness: Evidence for implicit control. *Consciousness and Cognition: An International Journal*, *16*(2), 421–435. https://doi.org/10.1016/j.concog.2006.06.010

Schmidt, J. R., & De Houwer, J. (2012a). Contingency learning with evaluative stimuli: Testing the generality of contingency learning in a performance paradigm. *Experimental Psychology*, *59*(4), 175–182. https://doi.org/10.1027/1618-3169/a000141

Schmidt, J. R., & De Houwer, J. (2012b). Learning, awareness, and instruction: Subjective contingency awareness does matter in the colour-word contingency learning paradigm. *Consciousness and Cognition: An International Journal*, *21*(4), 1754–1768. 10.1016/j.concog.2012.10.006

Schmidt, J. R., De Houwer, J., & Rothermund, K. (2016). The Parallel Episodic Processing (PEP) model 2.0: A single computational model of stimulus-response binding, contingency learning, power curves, and mixing costs. *Cognitive Psychology*, *91*, 82–108. https://doi.org/10.1016/j.cogpsych.2016.10.004

Schmidt, J. R., Giesen, C. G. & Rothermund, K. (2020). Contingency learning as binding? Testing an exemplar view of the colour-word contingency learning effect. *Quarterly Journal of Experimental Psychology*(73(5)), 739–761.

Shin, Y. K., Proctor, R. W., & Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychological Bulletin*, *136*(6), 943–974. https://doi.org/10.1037/a0020541

Stahl, C., & Unkelbach, C. (2009). Evaluative learning with single versus multiple unconditioned stimuli: The role of contingency awareness. *Journal of Experimental Psychology. Animal Behavior Processes*, *35*(2), 286–291. https://doi.org/10.1037/a0013255

Stahl, C., Unkelbach, C., & Corneille, O. (2009). On the respective contributions of awareness of unconditioned stimulus valence and unconditioned stimulus identity in attitude formation through evaluative conditioning. *Journal of Personality and Social Psychology*, *97*(3), 404–420. https://doi.org/10.1037/a0016196

Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.

Xu, G., & Mordkoff, J. T. (2020). Reliable correlational cuing while controlling for most-recent-pairing effects. *Frontiers in Psychology, 11:592377*. https://doi.org/10.3389/fpsyg.2020.592377

Table 1

German (English translation) target words, grouped by valence.

Positive		Negative	
Geschenk (gift)	Leben (life)	Bomben (bombs)	Krieg (war)
Traum (dream)	Musik (music)	Drogen (drugs)	Mord (murder)
Frieden (peace)	Jubel (joy)	Hinrichtung (execution)	Tumor (tumor)
Geburt (birth)	Umarmung (hug)	Gewehre (guns)	Unfall (accident)
Blumen (flowers)	Liebe (love)	Krankheit (sickness)	Verbrechen (crime)
Freund (friend)	Urlaub (holiday)	Krebs (cancer)	Virus (virus)

Table 2 Linear mixed-effects model analyses of trial RT (nested within participants).

	Model 1			**Model 2**			**Model 3**			**Model 4**		
Predictors	*b*	*SE*	*p*	*b*	*SE*	*p*	*b*	*SE*	*p*	*b*	*SE*	*p*
(Intercept)	-1.85	0.02	**<0.001**	-1.85	0.02	**<0.001**	-1.85	0.02	**<0.001**	-1.85	0.02	**<0.001**
Contingency (C)	0.03	0.00	**<0.001**	0.01	0.00	0.075	0.01	0.00	0.070	0.01	0.00	0.202
Previous response[a] (PR)				0.04	0.00	**<0.001**	0.04	0.00	**<0.001**	0.04	0.00	**<0.001**
Distance (D)							0.02	0.00	**<0.001**	0.02	0.00	**<0.001**
PR×D							-0.05	0.00	**<0.001**	-0.05	0.00	**<0.001**
C×Group[b]										0.03	0.01	**0.010**
Random Effects												
σ^2	0.09			0.09			0.09			0.09		
τ_{00}	0.06 $_{ID}$			0.06 $_{ID}$			0.06 $_{ID}$			0.06 $_{ID}$		
ICC	0.39			0.39			0.39			0.39		
N	129 $_{ID}$			129 $_{ID}$			129 $_{ID}$			129 $_{ID}$		
Observations	52766			52766			52766			52766		
Marginal R^2 / Conditional R^2	0.001 / 0.393			0.003 / 0.394			0.007 / 0.398			0.013 / 0.399		
AIC	24033.150			23889.775			23545.186			23628.241		

Note. Only effects of theoretical relevance are presented. In Model 4, all other non-displayed effects were not significant. [a]*Previous response* refers to the response relation (same vs. different) between the current response and the response on the last previous occurrence of the same nonword. [b]*Group* refers to incidental vs. instructed learning group, manipulated between subjects.

Figure 1

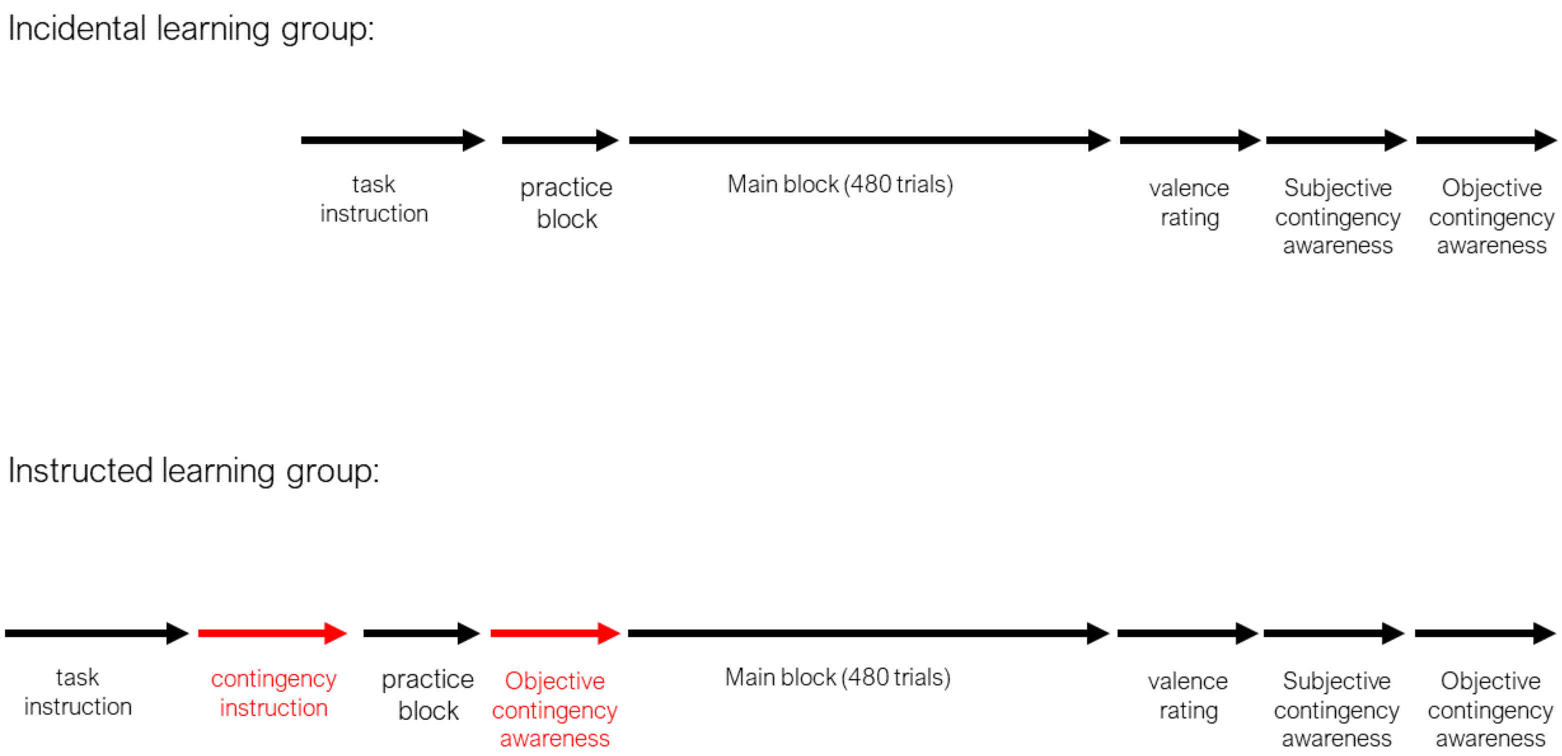

Figure 1. *Sequence of the experimental procedure in the incidental and instructed learning group. The experimental procedure of the incidental learning group is identical to the experimental procedure in Schmidt and De Houwer (2012a, Exp. 2). For the instructed learning group, additions were made to the procedure (highlighted in red; see main text for details).*

Figure 2

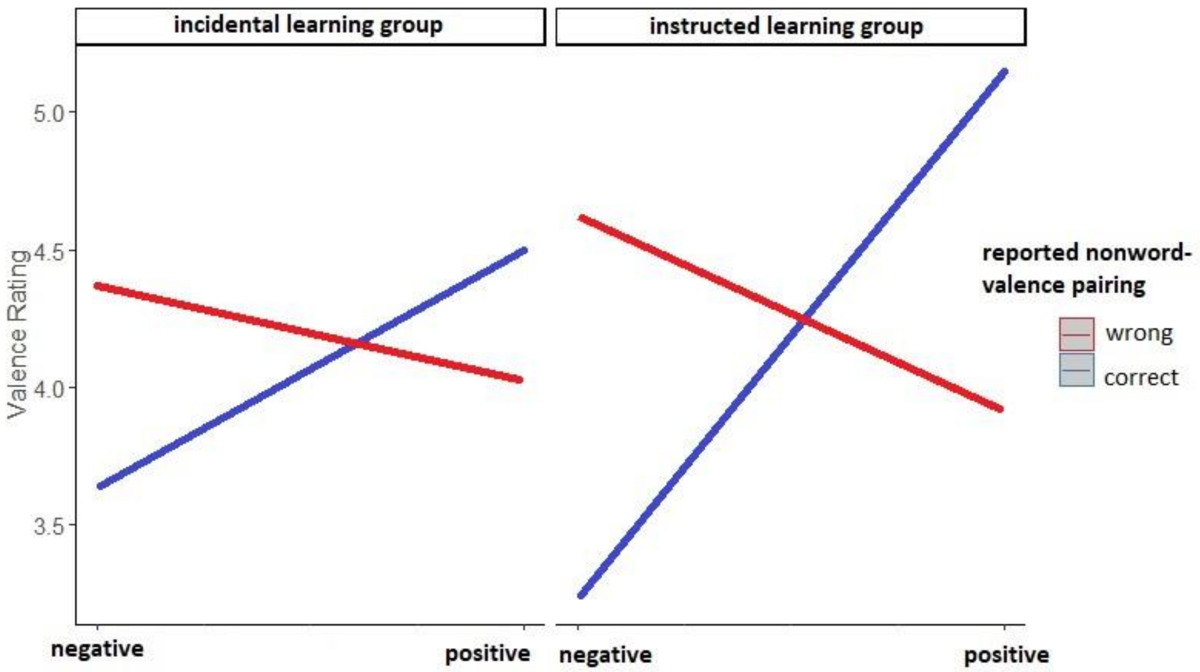

Figure 2. *Mean valence ratings for nonwords as a function of high contingency nonword-valence pairing, group, and reported nonword-valence pairing (correct responses indicate item-specific contingency awareness, wrong responses indicate absence of item-specific contingency awareness). EC effects are indicated by more positive (negative) ratings for nonwords paired with positive (negative) valence during high-contingency pairings. Note that EC effects are only present when item-specific contingency awareness existed (blue lines), but were even slightly reversed when item-specific contingency awareness was absent (red lines).*

Appendix

Additional Analyses

This section contains results with regard to standard analyses of CL effects (i.e., without controlling for effects of previous response effects) in RTs, error rates, and evaluative ratings, as well as their relations to measures of contingency awareness (subjective and objective). As the incidental learning group of the present study is a direct replication of Experiment 2 from Schmidt and De Houwer (2012a), we will first present analyses that were also performed by the authors to allow for easy comparison of findings. To anticipate, the present results closely replicate findings from Schmidt and De Houwer (2012a) with regard to contingency learning effects (RT and error rates) as well as with regard to EC effects (valence ratings for nonwords) for the incidental learning group.

Response Latencies

Mean RT in the valence contingency learning task were entered into a 2 (Contingency) × 2 (Valence) × 2 (Group) mixed factors Analysis of Variance (ANOVA). This yielded a significant main effect of contingency, $F(1,127)=48.06$, $p<.001$, $\eta_p^2=.27$, due to faster performance on high (565 ms) compared with low contingency trials (574 ms). This effect was qualified by a significant Contingency × Group interaction, $F(1,127)=8.02$, $p=.005$, $\eta_p^2=.06$ (Figure A1). Follow-up t-tests showed that CL effects (computed as $\Delta=RT_{low}-RT_{high}$) were larger for the instructed group than for the incidental learning group. CL effects significantly differed from zero in both groups (instructed learning: $\Delta=13$ ms; $t[64]=5.81$, $p<.001$, $d_z=1.02$ vs. incidental learning: $\Delta=6$ ms, $t[63]=3.96$, $p<.001$, $d_z=0.70$; CL effect for RT in Experiment 2 of Schmidt & De Houwer, 2012b,: $d_z=.90$). No other effect was significant.

Error rates

A 2 (Contingency) × 2 (Valence) × 2 (Group) mixed factor ANOVA on error rates yielded a significant main effect of contingency, $F(1,127)=14.56$, $p<.001$, $\eta_p^2=.10$, reflecting fewer errors on high (4.3%) compared with low (5.7%) contingency trials. This effect was qualified by a significant Group × Contingency interaction, $F(1,127)=5.01$, $p=.027$, $\eta_p^2=.04$. Follow-up t-tests revealed that this interaction was again due to larger CL effects in the instructed group ($\Delta=2.2\%$, which was significantly different from zero, $t[64]=3.40$, $p=.001$, $d_z=0.60$), than in the incidental learning group, ($\Delta=0.6\%$, which was not significantly different from zero, $t[63]=1.68$, $p=.097$, $d_z=0.30$; CL effect for errors in Experiment 2 of Schmidt & De Houwer, 2012b: $d_z=.46$). No other effect was significant.

Evaluative Conditioning Effects

We analyzed explicit evaluative ratings for nonwords with a 2 (Valence) × 2 (Group) mixed factors ANOVA. This yielded a main effect of valence, $F(1,127)=33.52$, $p<.001$, $\eta_p^2=.21$, indicating that nonwords which were predictive for positive targets in the valence contingency learning task were evaluated more positively ($M=4.7$) than nonwords which were predictive for negative targets ($M=3.7$). The Valence × Group interaction was significant as well, $F(1,127)=7.53$, $p=.007$, $\eta_p^2=.06$. Follow-up paired t-tests showed that EC effects (computed as the difference between ratings of positive-predictive minus negative-predictive nonwords) were stronger for the instructed learning group, and differed significantly from zero, $t(64)=5.34$, $p<.001$, $d_z=0.94$ (Figure A2), but were still significant even in the incidental learning group, $t(63)=2.52$, $p=.014$, $d_z=0.45$, replicating EC effects in Experiment 2 of Schmidt and De Houwer (2012b; $d_z=0.60$).

Contingency Awareness[3]

Objective awareness after main experiment. As objective awareness measure, participants had to report whether they felt a given nonword was mostly paired with positive or negative target words by pressing P or N, respectively. The accuracy of these responses was

[3] Contingency awareness analyses were conducted in a similar fashion as was done in Schmidt & De Houwer (2012a) to allow for an easier comparison of findings. We want to point out, however, that post-hoc classifications of participants into aware vs. unaware participants faces many methodological problems, as for instance regression to the mean (see Shanks, 2017, for details). Therefore, results of these analyses should be interpreted with caution.

coded for each nonword and averaged per person. Then, mean accuracy scores for each group we computed. Mean accuracy scores for the second objective awareness test differed significantly between groups, indicating higher accuracy (73.8%) for the instructed learning group compared with the incidental learning group (60.5%), $t(127)=3.05$, $p=.003$, $d=0.54$. However, note that accuracy scores were significantly better than chance (50%) for both, the instructed learning group, $t(64)=7.21$, $p<.001$, $d_z=1.26$, as well as for the incidental learning group, $t(63)=3.73$, $p<.001$, $d_z= 0.66$; Schmidt and De Houwer (2012b) obtained comparable frequencies (59.6% accuracy), which did not differ from chance, though.

Subjective awareness. Overall, 41 (32%) of the participants reported noticing contingencies by giving a yes or no response (subjective awareness). The number of subjectively aware participants differed significantly between groups (instructed learning: 43%, incidental learning: 20%), $t(127)=2.84$, $p=.005$, $d=0.50$. In a follow-up test, we assessed objective awareness after the experiment as a function of subjective awareness. Participants who were subjectively aware of contingencies had higher accuracy on the objective awareness measure (81.1%) than subjectively unaware participants (60.8%), $t(127)= 4.51$, $p<.001$, $d=0.79$.

Correlations

Table A1 presents correlations between all effects across and within groups. As can be seen, nearly all effects correlated significantly and positively with each other in the overall sample. Importantly, intercorrelations within groups indicated that this pattern was driven by the instructed learning group, whereas all effects were uncorrelated in the incidental learning group.

Table A1

Means, SD and intercorrelations of all effects across and within groups

	Variable	*M*	*SD*	1	2	3	4	5
Overall	1. Group	-	-					
	2. CL_{RT}	9.50	15.78	.24**				
	3. CL_{Err}	0.01	0.04	.19*	.40**			
	4. EC effect	0.96	1.93	.24**	.41**	.46**		
	5. obj. Awareness	0.67	0.26	.26**	.18*	.31**	.42**	
	6. subj. Awareness	0.32	0.47	.24**	.27**	.12	.24**	.37**
Group	Variable	*M*	*SD*		2	3	4	5
Instructed learning	2. CL_{RT}	13.28	18.44					
	3. CL_{Err}	0.02	0.05		.48**			
	4. EC effect	1.41	2.12		.58**	.57**		
	5. obj. Awareness	0.74	0.27		.21	.35**	.51**	
	6. subj. Awareness	0.43	0.50		.35**	.11	.33**	.45**
Incidental learning	2. CL_{RT}	5.66	11.42					
	3. CL_{Err}	0.01	0.03		.04			
	4. EC effect	0.50	1.59		-.10	.11		
	5. obj. Awareness	0.61	0.23		-.03	.13	.19	
	6. subj. Awareness	0.20	0.41		-.02	-.01	-.05	.15

Note. *M* and *SD* are used to represent mean and standard deviation, respectively. Group = experimental manipulation of instructed (coded as 1) vs. incidental learning (coded as 0). CL = contingency learning effect (see Table 2 for effect computation). EC= Evaluative Conditioning effects. Obj. Awareness=accuracy in objective awareness test at the end of the study. Subj. Awareness= subjective awareness rating (yes=1, no=0). * indicates $p < .05$. ** indicates $p < .01$.

Figure A1

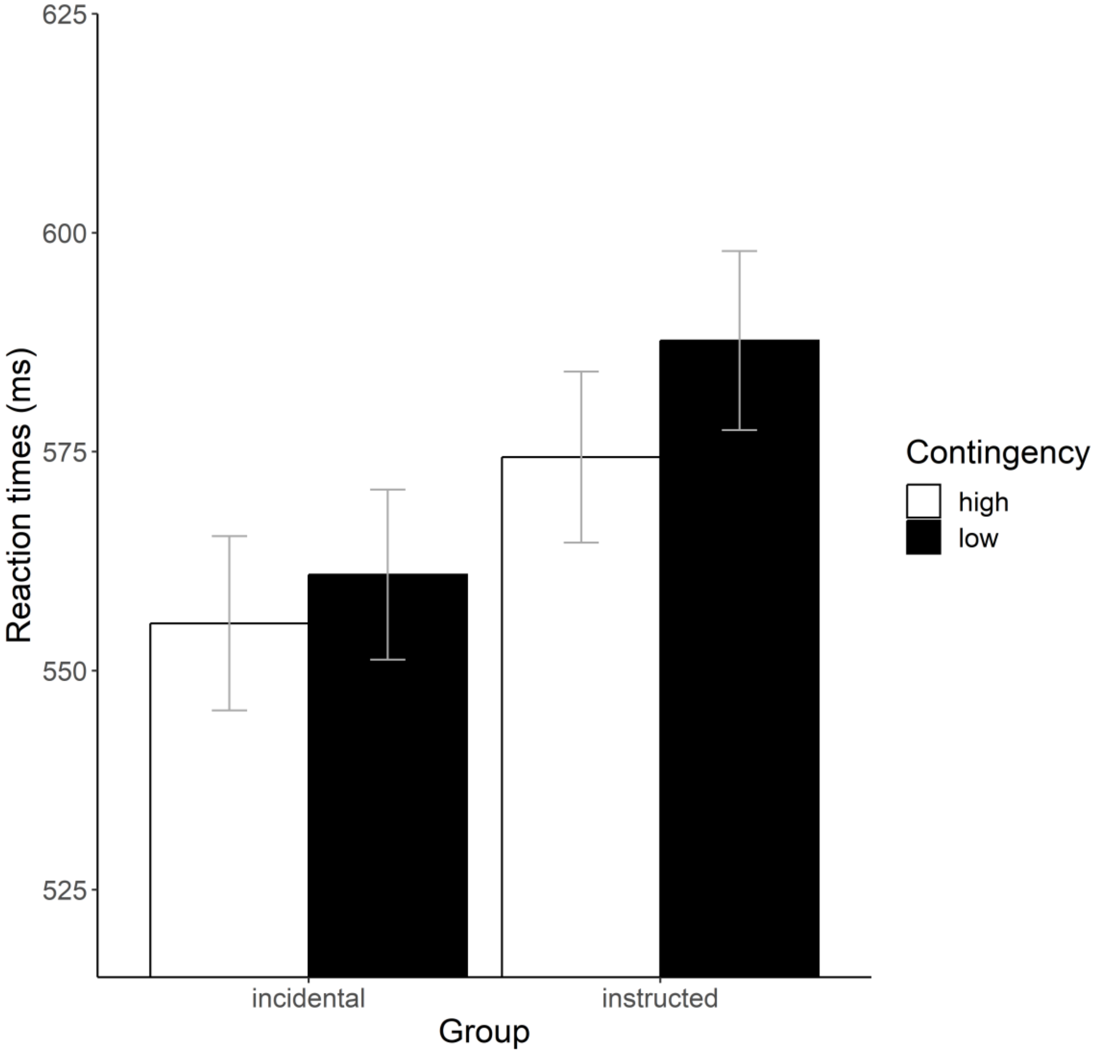

Figure A1. *Mean response latencies as a function of contingency and group. Error bars reflect standard error of the mean.*

Figure A2

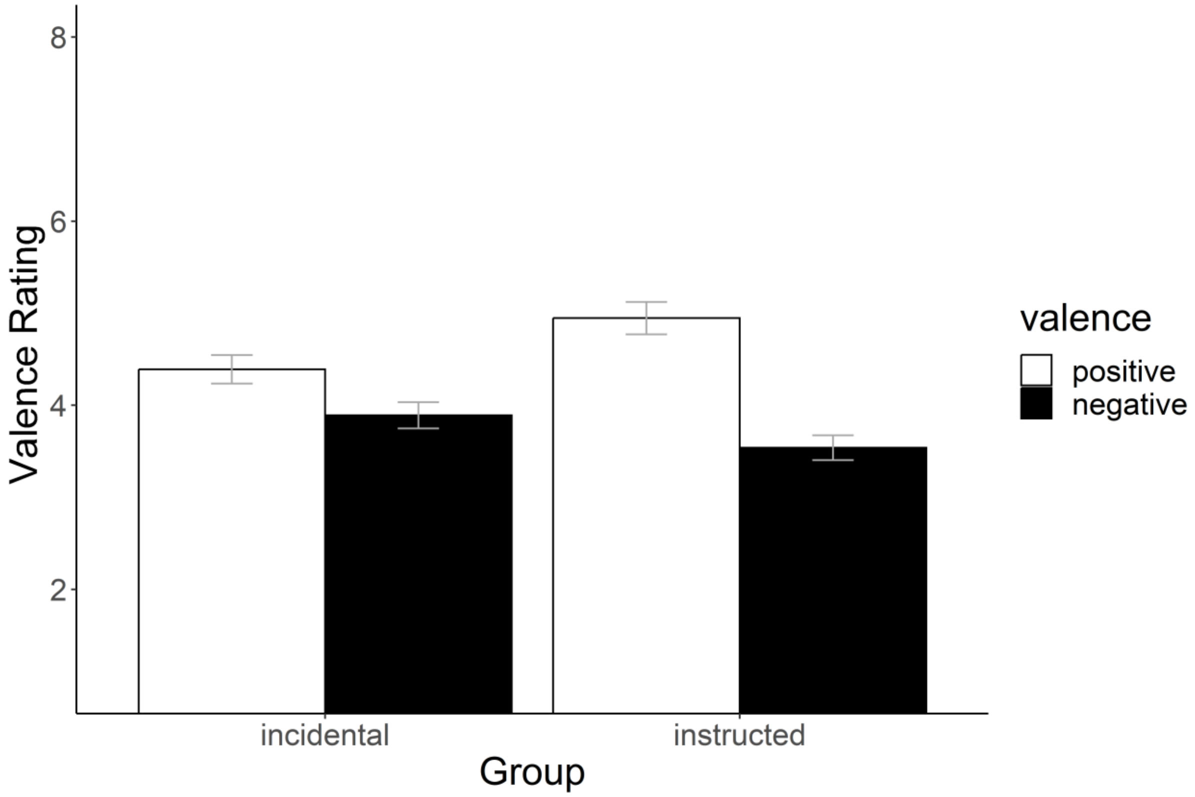

Figure A2. Mean valence ratings for nonwords as a function of Group and Valence of high contingent nonword-target pairings in the valence contingency learning task. Error bars reflect standard error of the mean.