A Dual-Force Perspective on Evaluative Conditioning

Inauguraldissertation

zur Erlangung des Doktorgrades

der Humanwissenschaftlichen Fakultät

der Universität zu Köln

nach der Promotionsordnung vom 18.12.2018

vorgelegt von

Lea Mareike Sperlich

aus Speyer

Februar 2024

"For there is nothing either good or bad, but thinking makes it so."

— William Shakespeare

Erstgutachter: Prof. Dr. Christian Unkelbach

Zweitgutachter: Prof. Dr. Hans Alves

Abstract

People constantly attribute positive or negative qualities to objects in their environment. What the organism considers to be positive or negative depends on the context in which it currently finds itself. While objects may be perceived as positive in one context, they may appear negative in another context. The social environment in which we operate is generally more positive than negative. In many cases, this positive environment results in negative information having advantages over positive information (i.e., a so-called 'negativity bias'). For example, negative information tends to be better remembered. Interestingly, there's hardly any evidence for this otherwise well-documented advantage of negative information in experiments on Evaluative Conditioning (EC). EC is defined as the change in liking of a conditioned stimulus (CS) due to its pairing with a positive or negative unconditioned stimulus (US). One reason for the absence of a negativity bias in EC experiments might be their symmetric structure, which does not reflect the information ecology of the real world. Other similar structural differences could also explain why typical EC experiments consistently elicit EC effects in the laboratory, while EC rarely occurs in the real world. In this thesis, I investigate *when* EC effects occur, focusing on the interplay between environment and organism. Additionally, I examine whether the difference in the structure of the real world and EC experiments can explain the absence of a negativity bias in EC experiments. The findings show that, in addition to the environment, the organism plays a crucial role in perceiving CS and US pairings. Furthermore, the thesis provides evidence that a cognitive-ecological approach can contribute to explaining the absence of a negativity bias in EC.

Zusammenfassung

Menschen schreiben Objekten in der Umwelt ständig positive oder negative Attribute zu. Was als positiv und was als negativ wahrgenommen wird, ist dabei auch abhängig vom Kontext, in dem sich der Organismus gerade befindet. Während Objekte in einem Kontext als positiv wahrgenommen werden, können sie in einem anderen Kontext negativ erscheinen. Die soziale Umwelt, in der wir uns bewegen, ist grundsätzlich mehr positiv als negativ. In vielen Fällen führt diese positive Umgebung dazu, dass negative Informationen einen Vorteil gegenüber positiven Informationen haben (d.h., zu einem sogenannten ‚negativity bias'). Solche negativen Informationen werden zum Beispiel besser erinnert. In Experimenten zu Evaluativer Konditionierung (EC) gibt es interessanterweise kaum Evidenz für diesen ansonsten gut belegten Vorteil für negative Information. EC wird definiert als die Veränderung der Präferenz eines konditionierten Stimulus (CS) aufgrund seiner Paarung mit einem positiven oder negativen unkonditionierten Stimulus (US). Ein Grund für das Fehlen eines negativity bias in EC-Experimenten könnte ihr symmetrischer Aufbau sein, der nicht der Informationsökologie der realen Welt entspricht. Andere ähnliche strukturelle Unterschiede könnten auch erklären, warum typische EC-Experimente im Labor konsistent EC-Effekte hervorrufen, während EC in der realen Welt eher selten auftritt. In dieser Arbeit untersuche ich deswegen, *wann* EC-Effekte entstehen, mit einem Fokus auf das Zusammenspiel von Umwelt und Organismus. Zusätzlich überprüfe ich, ob der Unterschied im Aufbau der realen Welt und EC-Experimenten das Fehlen eines negativity bias in EC-Experimenten erklären kann. Die Befunde zeigen, dass der Organismus, neben der Umwelt, einen wichtigen Faktor für die Wahrnehmung von CS und US als Paarung darstellt. Des Weiteren liefert diese Arbeit Evidenz dafür, dass ein kognitiv-ökologischer Ansatz zur Erklärung des fehlenden negativity bias in EC beitragen kann.

Preface

The thesis consists of a blend of cumulative work and a monograph. One paper has already been published, another is under review, and a third is currently in preparation.

Chapter 4.1 is based on the following article:

Sperlich, L. M., & Unkelbach, C. (2022). When do people learn likes and dislikes from co-occurrences? A dual-force perspective on evaluative conditioning. *Journal of Experimental Social Psychology, 103*, 104377. https://doi.org/10.1016/j.jesp.2022.104377

Christian Unkelbach and I collaborated on the development of the theoretical derivation. I programmed all experiments, implemented the data collection, and conducted all data analyses. We wrote the manuscript together.

Chapter 5.1 is based on the following manuscript:

Sperlich, L. M., & Unkelbach, C. (2024). *Gute Gestalt – why top-down input moderates EC, but bottom-up input doesn't.* Manuscript in preparation. Social Cognition Center Cologne, University of Cologne.

We developed the conceptual framework and all experimental designs together. I programmed all experiments, collected the data and conducted the analyses. I wrote the manuscript.

Chapter 6.1 is based on the following manuscript:

Sperlich, L. M., & Unkelbach, C. (2024). *Why is there no negativity bias in Evaluative Conditioning? – A cognitive-ecological answer.* Manuscript submitted to publication. Social Cognition Center Cologne, University of Cologne. (under review at the Journal of Personality and Social Psychology).

We developed the theoretical framework and all experiments together. I programmed all experiments, collected the data and conducted the analyses. We wrote the manuscript together. Wilhelm Hofmann provided us with the data for the re-analysis of meta-analytic data.

Some adjustments were made to the formatting and citation style to align with the dissertation's layout. However, the content of the articles and manuscripts remains unchanged.

Materials, preregistrations, pretests, data and analysis codes for the reported studies are available from Open Science Framework (OSF).

OSF link Chapter 4: https://osf.io/57ux2/?view_only=b5f058da184a4d24b2a951f44ef30ab9

OSF link Chapter 5: https://osf.io/2jzfk/?view_only=f85bbf1f539046bd93653c4f6756298c

OSF link Chapter 6: https://osf.io/9qbhr/?view_only=0167ad94e3a7437ca3cfab48d9358b3f

Diese Dissertation wurde von der Humanwissenschaftlichen Fakultät der Universität zu Köln im Juni 2024 angenommen.

Acknowledgments

First of all, I would like to thank my supervisor, Christian Unkelbach. Your trust in me and my work has led me to this point, where I am writing acknowledgments for my doctoral thesis (which I still can't quite believe). Over the past few years, your enthusiasm for research and science has inspired me, and I am more than grateful for that.

I would like to thank Hans Alves, who has proven that research isn't all boring – it's actually pretty interesting and relatable. Your creativity in coming up with experiments and your ability to explain complicated things in a simple way always impress me. As do your karaoke skills.

I want to thank everyone who's been there with me during my PhD journey. You've made coffee breaks into lively discussion sessions, kept me going with salad, pretzels, and snacks and always had an open ear for questions and problems. It wouldn't have been half as nice without you.

I would like to thank Moritz Ingendahl for his very valuable feedback regarding my dissertation. I would like to thank Tobias Wingen and Alexandra Fleischmann for giving me such a warm welcome at SoCCCo. I would like to thank Tabea Zorn, not only for being a great colleague and companion at conferences, but also for supporting conversations, both within and outside of work contexts.

I would like to thank Mareike Westfal, who not only put in an incredible amount of effort in proofreading my dissertation but also provided immense motivation beyond that. There was always a smile on my lips after your visits. Thank you for your great support, your motivational speeches and your little attentions that sweetened my day every time. I would like to thank Felix Speckmann for sharing his office with me and for letting me learn a lot from him. You are the master of good examples, an incredibly good lecturer, and one of the smartest people I know. Thank you for your consistent support, for the hugs, and for transforming our office into a little jungle. It is Wednesday, my dude.

I would like to thank Miriam Reiter for important suggestions and for her constant emotional support and helpful advice in all life situations. You add a lot to my life, and I'm glad you're here. Thank you for everything! I would like to thank my friends, especially Lisa Gödelmann, Mara Guntermann, Anna Gorthmanns, Julia Leng, Hannah Reuter, Caro Schüler, Vanessa Stösser, Leonie Stuthmann, and Vanessa Wagner. Thank you for your motivating words and actions when they were needed.

I would like to thank Frederike Völzer for the utmost support one could wish for. Thank you for always having my back, for motivating me to take breaks from my desk, for making me laugh when I needed it most, for always being there when I needed you. Fritzi, thank you for being in my life.

I would also like to thank my family, who made all this possible. Thank you for your constant support. Thank you, Heri & Ulla, for every visit to you being a little break. And thank you, Annette, for always being there to listen to me. A final thank you goes to Jona. I miss you.

List of Figures

List of Tables

CONTENT

General Introduction

What we like and what we dislike guides our behavior in a complex world (Walther & Langer, 2008; De Houwer & Hughes, 2020). However, not all of our preferences are pre-determined (e.g., preferences for sugar and fatty food; De Houwer & Hughes, 2020); instead, most preferences are shaped over the course of a lifespan (e.g., Baeyens et al., 1990; Rozin, 1982; Rozin & Millman, 1987), beginning as early as childhood (e.g., Doyle et al., 2022; Halbeisen et al., 2017; Richmond et al., 2017). For instance, one person might have a strong affinity for olives, while another might even dislike their scent. Similarly, there are varying degrees of enthusiasm in soccer, ranging from exuberant support for a particular club to complete indifference towards the sports as a whole. Another example are preferences in pets, with some individuals loving aggressive pit bulls, and others expressing hate for even the cutest dogs. Thus, individual preferences exhibit considerable variability. For that reason, researchers have examined different ways to modulate these preferences; with different researchers investigating different types of (evaluative) learning effects. One such evaluative learning effect is called Evaluative Conditioning (EC; De Houwer & Hughes, 2020). EC is defined as the change in the liking of a previously neutral stimulus (conditioned stimulus, CS) due to its pairing with a positive or negative stimulus (unconditioned stimulus, US). Thus, typically the CS is liked more when presented with a positive US and liked less when presented with a negative US (De Houwer, 2007; but see, e.g., Alves & Imhoff, 2023; Fiedler & Unkelbach, 2016). In our complex environment, we are constantly exposed to such neutral (CS) and valenced stimuli (US). Yet, EC effects are not constantly observed in everyday life. If they were, "sooner or later nearly all attitude objects would show the same indifferent evaluation" (Hütter et al., 2014, p. 633). In this thesis, I therefore attempt to answer the question of why EC is not observed more frequently in the real world, despite being considered a robust effect in psychological research (e.g., Fulcher & Hammerl, 2001; Moran, 2024; Walther, 2002). Thus, EC as a way of forming and changing attitudes[1] forms the core of this thesis.

[1] Based on Ajzen & Fishbein (2000), the term 'attitude' refers to the evaluation of a concept, object or behavior on dimensions like good-bad, pleasant-unpleasant or desirable-undesirable.

As EC is embedded in a broader field of learning psychology, in Chapter 1, I will give a definition of learning that allows to distinguish between different forms of learning (i.e., non-associative vs. associative). I will pay special attention to classical conditioning and operant conditioning, because both are related to EC. The former shares procedural similarities with EC, the latter is, like EC, an instance of evaluative learning (De Hower & Hughes, 2020). While there are other types of learning (e.g., learning through observation; Bandura, 1977), they are not within the scope of this thesis. Rather, the central focus is directed towards EC as the core of this thesis. Thus, I will describe EC in more detail in Chapter 2. In this chapter, I will focus is on the distinction of EC's structural properties (i.e., stimulus-driven bottom-up input) and influences of the organism on EC (i.e., top-down input). I will distinguish between process models of EC, which examine the underlying processes of EC (i.e., *how* EC occurs) and the dual-force model, that explores the interplay of bottom-up and top-down input (i.e., *when* EC occurs). Following this clarification, I will address the *Evaluative Information Ecology* (EvIE) in the social environment versus EC experiments in Chapter 3, with a focus on valence asymmetries. Chapter 4 and 5 present empirical tests of the dual-force model introduced in Chapter 2. In Chapter 6, predictions based on the EvIE model presented in Chapter 3 are empirically tested. Finally, the findings from previous chapters are discussed in Chapter 7.

Chapter 1: Definitions of Learning

Although learning as a field of research has a long tradition, there is little agreement on a precise definition of learning (De Houwer et al., 2013; De Houwer & Hughes, 2023). While some researchers do not define learning at all (e.g., Bouton, 2007; Schwartz et al., 2002), most textbooks describe learning as changes in the behavior that are due to experiences (Lachman, 1997). However, this simple functional definition of learning has its difficulties. For example, the assumption that learning is represented in changes in behavior is not compatible with the observation that previously learned behavior might not be shown in an immediate behavior change but only at a later point in time (e.g., Tolman & Honzik, 1930). To explain this occurrence of 'latent' (De Houwer et al., 2013) or 'behaviorally silent' learning (Dickinson, 1980), learning is described in alternative definitions via changes in the organism. These mechanistic definitions of learning assume that experience leads to a change in the organism that persists over time and might only eventually be expressed in behavior (e.g., Anderson, 2000). According to De Houwer and colleagues (2013), mechanistic definitions of learning describe learning either as “mechanisms” of behavior (Domjan, 2010, p. 17), as a process (Lachman, 1997), “some kind of internal changes” or “as involving a specific mental process” (De Houwer et al., 2013, p. 632). However, these definitions raise some difficulties (e.g., how to determine that changes in the organism have occurred).

1.1 Phylogenetic versus Ontogenetic Learning

De Houwer and colleagues (2013) thus propose a functional definition of learning, that aims to solve the problems of simple functional and mechanistic definitions. To do so, they distinguish between phylogenetic and ontogenetic learning. From an evolutionary perspective, organisms are in a state of constant development. In other words, they are constantly adapting to their environment. This adaptation takes place over generations. Traits that ensure survival are passed on to the next generation because they have an adaptive advantage. De Houwer and colleagues (2013) refer to this as phylogenetic adaptation. In contrast, the object of study in learning psychology is referred to as ontogenetic adaptation. Here, the adaptation of an organism to its environment is not considered over generations, but during its lifespan. Accordingly, learning is defined as "changes in the behavior of an organism that are the result of regularities in the environment of that organism" (De Houwer et al., 2013, p. 633). Although De Houwer and Hughes (2023) extended this definition (e.g., to include nonorganic entities), in the present work I rely on De Houwer et al.’s (2013) definition when referring to learning.

This more complex functional definition of learning facilitates distinguishing between different subclasses of learning:

(a) Non-associative learning

(b) Classical / Pavlovian conditioning

(c) Operant / Instrumental conditioning

I will describe these subclasses in more detail in the next sections. First, I delineate habituation and sensitization as variants of non-associative learning. In classical conditioning, I distinguish between S-S and S-R learning and present the Rescorla-Wagner model for contingency learning before elaborating on the concept of contiguity. Lastly, I introduce operant conditioning and examine the relevance of punishment and reward.

1.1.1 Non-Associative Learning

Non-associative learning "is considered a fundamental form of learning that can be observed across all animal phyla and most sensory modalities" (Schmid & Poon, 2012, p. 2475). In non-associative learning, regularities in the environment consist of only one stimulus that occurs repeatedly. Thus, there is no association or pairing with other environmental stimuli (Ioannou & Anastassiou-Hadjicharalambous, 2021). Early theories of non-associative learning distinguish two processes: *habituation* and *sensitization* (Groves & Thompson, 1970; Thompson, 2009).

Habituation. Thorpe (1956) suggested habituation to be the simplest form of learning (see also Hawkins, 1989; Rankin et al., 2009; Thompson & Glanzman, 2016). Habituation allows animals to disregard unimportant stimuli and concentrate selectively on significant ones (Thompson, 2009). By repeatedly presenting a stimulus, the intensity of the response to this stimulus decreases (De Houwer et al., 2013). For example, when hands are clapped out of a child's sight, this clapping initially leads to eye blinking. After repeated hand clapping, however, this reaction becomes weaker: the child habituates to the stimulus. However, after giving a sharp blow, the child will blink again. Humphrey (1933/2013) suggests that "blowing on the cradle requires a new adjustment on the part of the organism" (p. 142), a process that he called *dishabituation*. Thus, dishabituation occurs when introducing a new or novel stimulus causes the organism to regain responsiveness to the original stimulus (Groves & Thompson, 1970).

Sensitization. Yet, presenting an extraneous stimulus can even increase the reaction of the habituated stimulus (Thompson, 2009). This state of surpassing the baseline response of the habituated stimulus is called *sensitization* (Groves & Thompson, 1970). On a functional level, sensitization might alert organisms (Shettleworth, 2009). For example, when waiting for school to end, pupils might be more sensitive to the sound of a ringing bell (Sweatt, 2009).

1.1.2 Classical / Pavlovian Conditioning

In contrast to non-associative learning, classical conditioning requires the presence of at least two stimuli and the formation of relations between these stimuli (De Houwer et al., 2013). One of the best-known experiments on classical conditioning is the dog experiment by Ivan Pavlov (1927/2010). Pavlov observed that an unconditioned stimulus (UCS[2]; food) led to an autonomic response in dogs (UCR; salivation), while the mere presentation of a neutral stimulus, a bell, did not. However, after repeatedly presenting UCS and bell, the bell (conditioned stimulus; CS) alone triggered the conditioned response (CR; salivation). This response extinguishes after some time, a phenomenon that is called *extinction*: "Just as a CS–U[C]S pairing comes to evoke responding, subsequently presenting CS without the U[C]S on a number of trials eliminates […] that responding" (Bouton & Moody, 2004, p. 664). Extinction as well as other basic principles of conditioning (e.g., acquisition, spontaneous recovery, generalization; for an overview see Mazur & Odum, 2023) also apply to human conditioning (Koch & Stahl, 2017). One such paradigm to test classical conditioning in humans is the eyeblink paradigm (Mazur & Odum, 2023). In this task, a puff of air (UCS) leads to eyelid closure (UCR). When repeatedly paired with the UCS, a light or tone (CS) elicits the eyelid closure (CR), without the simultaneous presentation of the UCS (e.g., Clark & Squire, 1998; Clark et al., 2002; see also Lovibond & Shanks, 2002).

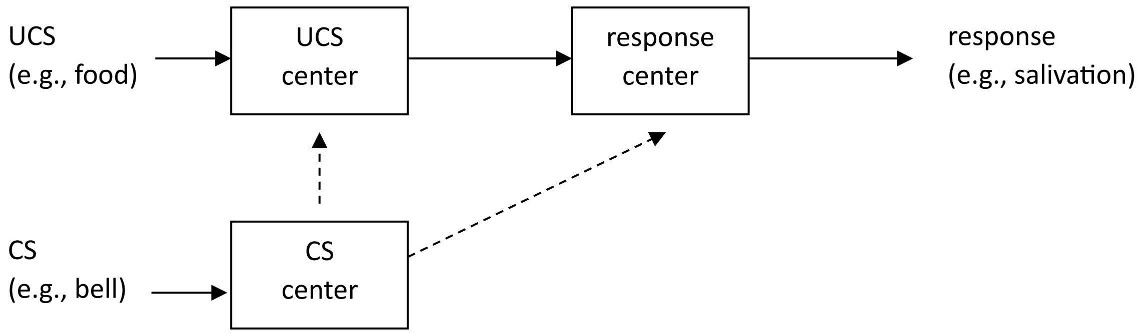
Stimulus-Response versus Stimulus-Stimulus Learning. Pavlov proposed stimulus substitution as a theory that might underlie classical conditioning. According to this theory, the CS substitutes the UCS. That is, the CS elicits the same response as the UCS after their joint presentation (Mazur & Odum, 2023). Pavlov assumed that, in the brain, the UCS presentation activates what Mazur and Odum (2023) call the *U[C]S center*. Similarly, the CS presentation activates a *CS center*,

[2] *Note.* In the literature, UCS and US are used rather synonymously, regardless of the type of conditioning. However, I will use the term UCS to refer to classical and operant conditioning and the term US to refer to EC to avoid confusion.

while the *response center*, when activated, produces the observed response via neural commands. The link between representations of UCS and UCR is considered innate, while links between CS and UCS (S-S) or CS and response (S-R) are thought to be formed during conditioning (Mazur & Odum, 2023; see Figure 1.1).

Figure 1.1

Schematic Illustration of Pavlov's Stimulus Substitution Theory.



Note. The rectangles represent the mental representation of UCS, CS and response. Mazur and Odum (2023) refer to these representations as U[C]S center, CS center and response center. The solid lines represent the assumed innate connection between UCS, UCS center and response. The dashed lines represent two types of associations that might form during conditioning (see also Koch & Stahl, 2017). Adapted from Mazur and Odum (2023).

S-R learning is in line with a behavioristic perspective, that focuses on observable behavior shaped by environmental events (e.g., Horowitz, 1992). S-S learning, on the other hand, requires some kind of mental representation. According to S-S learning, a link between CS and UCS is formed due to their joint presentation. Consequently, the CS can activate the mental representation of the UCS and thus elicit a response that is resembling the UCS' response. S-S learning thus incorporates cognitive elements, while S-R learning does not necessarily rely on such aspects (Mitchell et al., 2009).

Even though S-R learning might still play a role in learning (Dickinson, 1994; Pearce & Bouton, 2001; Rescorla, 1991), many conditioning phenomena (e.g., blocking, sensory preconditioning) are better explained by S-S learning (Mitchell et al., 2009). Compelling evidence supporting S-S learning is derived from UCS devaluation experiments (e.g., Rescorla, 1973). Rescorla investigated the role of CS-UCS-response links using a *conditioned suppression procedure*. In condition suppression procedures, animals (e.g., rats) initially learn a specific response (e.g., pressing a lever). Occasionally, a CS (e.g., a tone) occurs, followed by an aversive UCS (e.g., an electric shock). After a few trials, the animal decelerates or entirely stops showing the previously learned behavior when the CS is present (i.e., the rat *suppresses* this behavior). In the first phase of his experiments, Rescorla (1973) repeatedly presented a loud noise (UCS) with a light (CS). In the second phase, he habituated one group of rats to this sound (i.e., he presented the loud noise without the CS), while the other group did not receive such treatment (i.e., neither CS nor UCS was presented). If the S-R link was responsible for the response, the habituation procedure should not influence the CR. That is, no difference between the two groups of rats should be observed. If, however, associations were formed between CS and UCS (S-S link), the CR should be weaker in the experimental group (compared to the control group). As suppression was stronger in the experimental group, Rescorla's (1973) results support the second assumption. Findings from UCS inflation (i.e., increasing the UCS' intensity) provide further evidence for S-S learning (e.g., Rescorla, 1974; Sherman, 1978). Similar effects were found with human participants in fear conditioning (Davey & McKenna, 1983; White & Davey, 1989).

Contiguity. Another factor that was deemed important for classical conditioning is c*ontiguity* (i.e., spatial-temporal proximity) between CS and UCS. However, this assumption was challenged by several experimental findings (e.g., Koch & Stahl, 2017). One of these studies illustrates the so-called blocking effect (Kamin, 1968; 1969). In a classic study with rats, Kamin (1968) used a conditioned suppression procedure to demonstrate the effect. In the experimental group, he first presented a light (CS1) together with a shock (UCS). In a second step, the light (CS1) and a tone (CS2) were both presented with the UCS. In the control group, the first phase was skipped. Yet, both groups of rats received the exact same number of tone (CS2) and shock (UCS) presentations. Thus, the contiguity principle would predict no difference in the CR between experimental and control group. While the rats in the control group acted accordingly, the rats in the experimental group showed almost no CR at all. Kamin (1968) reasoned that the prior conditioning using the light somehow *blocked* the later conditioning of the tone. One simple explanation is that the rats in the experimental group had already learned that the light was a reliable predictor for the shock. Put differently, the tone did not provide any further information and thus, conditioning did not occur (Mazur & Odum, 2023).

Rescorla-Wagner model. These findings regarding the blocking phenomenon led to the development of the *Rescorla-Wagner model* (Mazur & Odum, 2023). The central assumption of this model is that learning will only occur, when the organism is surprised; that is, if the organism's expectations are violated. *What* the organism expects depends on what has been learned before. Thus, the model operates on a trial-by-trial basis (Mazur & Odum, 2023). The changes in associative strength accrued to a CS_A can be formalized as:

$$\Delta V_A = \alpha_A \times \beta \ (\lambda - \sum V_{all})$$

ΔV_A describes the change in the strength of the CS_i-UCS link. The α-value represents the learning rate parameters for CS_A (i.e., the salience of CS_A), the β-value represents these parameters for the UCS (i.e., the UCS intensity). These parameters are usually held constant across trials. Finally, λ "represents the asymptotic level of associative strength which each U[C]S will support" (Rescorla & Wagner, 1972, p. 76). In other words, the amount of learning in each trial (ΔV_A) depends on the difference between reality (i.e., the actual UCS strength: λ) and expectation (i.e., $\sum V_{all}$; that contains all stimuli that are present).

This model can explain Kamin's (1968; 1969) findings regarding the blocking effect. According to the Rescorla-Wagner model, the strongest association formation takes place in the first trial as the joint occurrence of CS and UCS is most surprising in this trial. In subsequent trials of joint CS_A-UCS presentations (assuming perfect contingency[3]), the effect of surprise decreases: the expectation term approaches the asymptote ($\lambda \approx \sum V_{all}$). Thus, if a second CS_X (that does not differ in salience from CS_A) is presented together with CS_A and the respective UCS, there's little or no potential left to be associated with the novel CS_X. Because there's no change of the UCS and because CS_A is part of CS_X' expectation term, the difference between reality (λ) and expectation ($\sum V_{all}$) remains small. Thus, changes in associative strength accrued to CS_X are either small or non-existent.

Even though the Rescorla-Wagner model is one of the most prominent theories in classical conditioning, there are some limitations to this model (e.g., CS preexposure effect; Mazur & Odum, 2023). Additionally, other theories can also explain observations like the blocking effect. For example, propositional theories[4] (e.g., Mitchell & Lovibond, 2002; De Houwer, 2009; Mitchell et al., 2009). These theories assume that during conditioning, propositions are formed that connect CS and UCS (e.g., "When I hear a bell, I will receive food", Lovibond & Shanks, 2002, p.186). Concerning the blocking effect, the formed propositions might be 'CS_A is always followed by the UCS', 'Adding CS_X does not change the outcome, thus CS_X does not add anything beyond CS_A', 'CS_A, and not CS_X, causes the UCS' (De Houwer et al., 2002; De Houwer, 2009).

1.1.3 Operant / Instrumental Conditioning

Similar to the S-S /S-R discussion in classical conditioning, representatives of operant conditioning speculated whether a response-stimulus (R-S) association is formed during conditioning, as assumed in early approaches. In operant conditioning, organisms learn to show (or to refrain from showing) a certain behavior by reinforcing (or punishing) the corresponding behavior (Koch & Stahl, 2017). In typical experiments, rats learn to press a lever (response), when pressing the lever is rewarded with food pellets (UCS). This behavior is instrumental, as it leads to consequences in the

[3] Contingency between CS and US can be formulated as $\Delta P = P(UCS|CS) - P(UCS|-CS)$: the difference ($\Delta$) between the probability (P) for the occurrence of the US following the CS and the probability of the UCS occurring without previous CS occurrence. Perfect contingency is present when $P(UCS|CS) = 1$ and $P(UCS|-CS) = 0$ ($\Delta P = 1$).

[4] The propositional approach can of course not be applied to invertebrates that do not have conscious beliefs (e.g., Aplysia; Mitchell et al., 2009).

environment that would otherwise not occur. Thus, in operant conditioning (as well as in classical conditioning), the focus lies on changing a behavior (Koch & Stahl, 2017).

Reinforcement and Punishment. To systematically investigate this environmental influence on behavior, Skinner (1938/2019) used an operant chamber (i.e., a Skinner box). Skinner defined a reinforcer on an operational level as a stimulus that changes the occurrence probability of a response. This definition implies that reinforcement does not necessarily have to be positive. A change in the occurrence probability of a response can also be induced by removing a negative stimulus (i.e., negative reinforcement). A third factor that can elicit changes in behavior is punishment. Punishment involves either adding an unpleasant stimulus or removing a pleasant stimulus (see Table 1.1).

Table 1.1

Reinforcement and Punishment in Operant Conditioning.

Consequence of behavior	Quality of stimulus	
	Pleasant	**Unpleasant**
Presenting stimulus	Positive reinforcement	Punishment
Removing stimulus	Punishment	Negative reinforcement

Note. Without consequences (i.e., without reinforcing or punishing the behavior), performance of the learned behavior will stop (i.e., extinction of behavior). Adapted from Koch & Stahl (2017).

Using the Skinner box, rats might for example be reinforced by feeding them pellets (positive reinforcement) or removing an electric shock (negative reinforcement). Similarly, they can be punished by removing food (i.e., removing a positive stimulus) or by giving an electric shock (i.e., adding a negative stimulus). Learning takes place as a consequence of these environmental changes. As soon as reinforcement or punishment discontinues, the learned behavior vanishes. That is, extinction occurs (cf. Chapter 1.1.2).

1.2 Evaluative Learning

Operant conditioning is a variant of evaluative learning. Based on De Houwer and colleagues' (2013) definition, the concept of environmental regularities allows one to distinguish between three types of such regularities: (a) regularities when a single stimulus is present, (b) regularities when two stimuli are present, and (c) regularities when behavior and stimuli are present (De Houwer & Hughes, 2020). Regularities in the presence of one stimulus are related to the mere-exposure effect (Zajonc, 1968); regularities in the presence of two stimuli to EC. Regularity learning in the presence of stimuli and behavior is related to the just mentioned operant conditioning.

1.2.1 Mere Exposure

One way to form attitudes towards a stimulus is to merely present the stimulus repeatedly. After such repeated presentation, the stimulus becomes more positive, an observation that is coined *mere-exposure effect* (Zajonc, 1968). The mere-exposure effect can for example explain why we might be quite neutral towards a song when we listen to it for the first time but start to like it after hearing it the third time. The effect also works in the opposite direction: When we did not like a song after hearing it for the first time, we might dislike it even more after hearing it again (cf. Brickman et al., 1972).

1.3 Outlook

However, attitude acquisition is not confined to the presence of one stimulus alone. As mentioned before, attitudes can also be shaped by presenting two stimuli together. For example, watching happy people (US) drinking an unfamiliar beverage (CS) might increase the liking for the respective beverage. In other words, the attitude towards the beverage becomes positive when paired with happy people. In the remainder of this thesis, I will focus on this type of evaluative learning, namely *Evaluative Conditioning (EC)*.

To introduce the core topic of this thesis, I will first define EC in more detail and give a brief overview of its moderators (Chapter 2). In this context, I will also distinguish the *how* question of EC from the *when* question. The former investigates "what cognitive processes intervene between mere co-occurrence and attitude formation or change" (Jones et al., 2010, p. 205), the latter examines the influence of bottom-up and top-down input on the encoding of CS and US into the same mental episode and is experimentally tested in Chapter 4 and 5. I will then review research on the *Evaluative Information Ecology (EvIE) model* and outline its implications for EC experiments (Chapter 3). I will give an overview of research on valence asymmetries and present an approach that might explain the lack of valence asymmetries in EC experiments. Chapter 6 then presents an empirical test of the EvIE approach presented in Chapter 3. Finally, I will discuss implications and limitations of the presented findings in the General Discussion in Chapter 7.

Chapter 2: Evaluative Conditioning

A first attempt to examine attitude formation in a well-controlled experiment was made by Staats and Staats (1958). The authors used valenced words (e.g., pretty) as USs and nationalities (e.g., Dutch) as CSs. Using pictures instead of words, Levey and Martin (1975) revisited this early work, calling it Evaluative Conditioning (EC; Martin & Levey, 1978). In their experiment, today referred to as *picture-picture paradigm*, participants sorted postcards of paintings and photographs into three categories: neutral, liked, or disliked (Levey & Martin, 1975). The experimenter then selected pictures from each category (while care was taken that the pictures did not differ in terms of complexity and discriminability). Liked and disliked pictures were used as USs, and neutral pictures as CSs. This procedure resulted in five pairs for each participant: CS + US_{pos}, CS + US_{neg}, CS + CS, US_{pos} + CS, US_{neg} + CS ($_{pos}$ for positive, $_{neg}$ for negative). After that, the picture pairs were presented to participants. Subsequent liking ratings showed that participants liked CSs presented with USs_{pos} better than the CSs presented with USs_{neg}. Since these early demonstrations, EC has become one of the most-used paradigms to study the acquisition of attitudes (Hütter, 2022). EC is deemed a robust effect (for a meta-analysis, see Hofmann et al., 2010; but see also Jones et al., 2010; De Houwer et al., 2005) that is studied in many different areas (for reviews see De Houwer et al., 2001; Gast et al., 2012; Walther et al., 2005), including learning psychology (e.g., Martin & Levey, 1978), social psychology (e.g., Olson

& Fazio, 2001, 2006; Walther, 2002), consumer research (e.g., Allen & Janiszewski, 1989; Ingendahl, Vogel et al., 2023; Sweldens et al., 2010), personality psychology (e.g., Ingendahl, Woitzel & Alves, 2023; Ingendahl & Vogel, 2023; Vogel et al., 2019), clinical psychology (e.g., De Houwer, 2020; Franklin et al., 2017; Glashouwer et al., 2018), and neuroscience (e.g., Bosshard et al., 2019; Coppens et al., 2006; Kuchinke et al., 2015).

As EC is similar to classical conditioning (Hütter, 2022), researchers have long disagreed as to whether EC is merely a sub-form of classical conditioning or whether it qualifies as a distinct form of learning (e.g., Baeyens & De Houwer, 1995; Blechert et al., 2008; Davey, 1994; Lipp & Purkis, 2005; Martin & Levey, 1994; see also Walther & Langer, 2010). This debate on EC as distinct phenomenon mainly focused on EC's independence of contingency awareness (i.e., awareness of the CS-US pairings) and its resistance to extinction. Both assumptions have been challenged lately (for a meta-analysis, see Hofmann et al., 2010; for a review see Sweldens et al., 2014; Moran et al., 2023). In this thesis, however, I will take one step back: Instead of asking *how* EC emerges, a question that is concerned with the role of contingency awareness in EC, I will ask *when* EC occurs, which is concerned with the interplay of organism and environment in EC regarding the construction of mental episodes.

Although EC can be defined in terms of a procedure or a theoretical process, De Houwer (2007) notes that EC is best defined as an effect. In line with this recommendation, I will refer to EC as a change in liking, that is elicited by the pairing of the CS with a positive or negative US. Defining EC "in terms of elements in the environment (stimulus pairings) without referring to mental processes and representations (e.g., formation of associative links; acquisition of propositional knowledge)" (Gast et al., 2012, p. 80) has the decisive advantage that it does not confine the effect of stimulus pairings on EC to a priori assumptions about mental processes. Put differently, this functional definition of EC as an effect (rather than a procedure or theory) allows different processes to take place without requiring assumptions about them. Thus, this definition allows one to ask *when* EC occurs without warranting an answer to the *how* question.

In the next section, I will first describe the typical structure of an EC experiment (i.e., the EC experiment's environment), and relate these structural elements to the experiments presented in Chapters 4-6. I will then elaborate on the role of the organism in previous EC research. Finally, I will address the *when* question of EC, and integrate environment and organism into a dual-force perspective.

2.1 The Environment

In a typical EC experiment, a CS (e.g., a picture of a Chinese ideograph) is repeatedly presented together with a US (e.g., a picture of a puppy). The US can either be (subjectively) positive or negative, whereas the CS is initially neutral. Through the repeated presentation (i.e., the pairing) of CS and US, the liking of the CS changes in direction of the respective US (De Houwer, 2007).

A basic EC experiment consists of three phases: (a) pre-acquisition phase, (b) acquisition phase, and (c) evaluation phase. The pre-acquisition phase is concerned with the choice of the CS. CSs can either be taken from pretested databases (e.g., IAPS: International Affective Picture System, Lang et al., 2008), selected based on pre-ratings of another sample, or based on pre-ratings of the current experimental sample (Moran et al., 2023). Regarding the acquisition phase, different types of trials with different stimulus combinations are possible. However, in this thesis, only one type of trial is used: CS_{neg} + US_{neg} and CS_{pos} + US_{pos} (cf. Moran et al., 2023). Because CSs are usually neutral in valence (Moran et al., 2023), the CS_{pos} does not refer to an initial positivity of the CS, but to the US it is presented with (i.e., a positive US: US_{pos}). The same applies to CS_{neg}. In the evaluation phase, changes in the CS can be captured using direct measures (e.g., "'how much do you like the CS'?", Moran et al., 2023, p. 249) or indirect measures (e.g., Implicit Association Test; cf. Mitchell et al., 2003). Directly asking participants is deemed to be a straightforward way to gain insights into a person's thoughts and feelings. In contrast, indirect measures might avoid socially desirable responses (Gawronski & Hahn, 2018). Yet, all experiments presented in this thesis used direct measures, asking participants to rate CS likings on Likert scales whit ranges varying between experiments.

2.1.1 The Experimental Structure

Typically, EC experiments consist of one or more trials. In each trial, a CS, a US, or a CS-US pairing is presented. This pairing of CS and US usually results in the CS acquiring a valence similar to the US (i.e., an EC effect; Moran et al., 2023). Usually, the stimuli or stimulus pairings are presented repeatedly. Even though EC effects were found in experiments using only one CS-US pairing[5] (Stuart et al., 1987), most experiments use more than one pairing presentation. However, evidence regarding the relation between stimulus co-occurrences and EC effects is inconclusive (Kurdi & Banaji, 2019; for a meta-analysis see Hofmann et al., 2010; for reviews see De Houwer, 2001; Moran et al., 2023). Some studies found that EC increases with the number of total CS-US pairings (e.g., Sachs, 1975; Stuart et al., 1987). For example, Baeyens and colleagues (1992) found an increase in EC effects for up to 60 CS-US pairings, whereas Bar-Anan and colleagues (2010) report equally strong EC effects for 24 and 64 pairings, and a smaller effect for 44 pairings; however, this difference was not significant. Yet, in a second Experiment, the authors did not find evidence for EC effects for less than 12 CS-US pairings. Staats & Staats (1959) found an EC effect with 48 CS-US co-occurrences but suggested that an increase of up to 108 CS-US pairings will produce the greatest extent of conditioning. In a similar study, the authors found an EC effect using 48 pairings (Staats et al., 1959). Levey and Martin (1987) argued the number of stimulus pairings to be inconsequential; however, they did not provide statistical evidence for their assumption (Kurdi & Banaji, 2019). Others found EC effects using only four CS-US pairings (Kurdi & Banaji, 2019), or as mentioned before, for even a single pairing (Stuart et al., 1987). In this thesis, CSs and USs co-occurred repeatedly; however, the total number of co-occurrences differs between experiments (with a minimum of 24 CS-US pairings).

However, the more interesting factor regarding this thesis is the ratio of positive and negative pairings presented in a typical EC experiment rather than the absolute number of pairings. Usually, this ratio is 1:1. Thus, in a hypothetical study that presents eight CS-US pairs, four of these pairings

[5] I will report the number of total CS-US pairings instead of trials, because the number of CS-US pairings within trials differs between the studies.

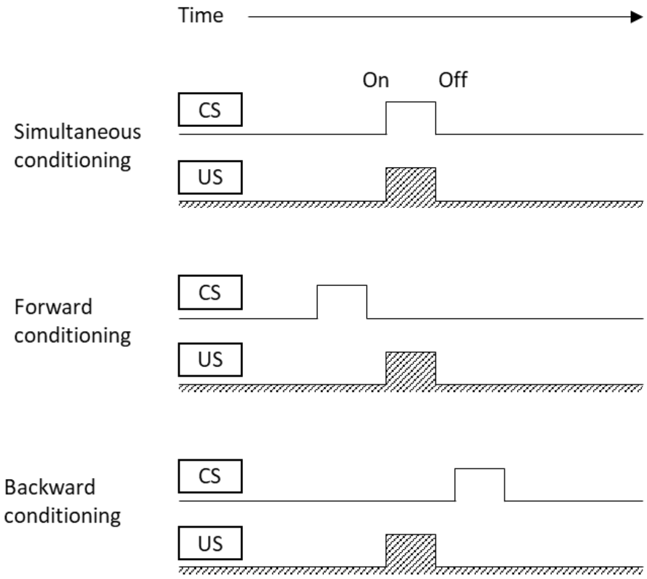
are positive (i.e., CS + US_{pos}) and four pairings are negative (i.e., CS + US_{neg}). The significance of this experimental structure will be further evaluated in Chapter 6, where we changed the ratio to 4:1 (i.e., presenting four CS-US_{pos} pairings, and one CS-US_{neg} pairing) or 1:4 (i.e., presenting one CS + US_{neg} pairing, and four CS-US_{pos} pairings), dependent on condition.

Additionally, pairings in EC can consist of *one-to-one* assignments or *one-to many* assignments (Stahl & Unkelbach, 2009). In one-to-one assignments, one CS is assigned to one US. In one-to-many assignments, one CS is assigned to various USs of the same valence. Is this thesis, only one-to-one assignments are used. EC effects were shown to be of similar strength for odor, taste, auditory, sensical verbal and visual USs, but most pronounced for electrocutaneous stimulation (Hofmann et al., 2010). However, EC studies usually use different liked and disliked USs rather than (mild) electric shocks (e.g., De Houwer et al., 2001; Glaser & Walther, 2012; Hughes et al., 2016). CSs can be of the same modality like USs (e.g., picture US – picture CS) or from a different modality than the US (e.g., sound US – picture CS). Experiments presented in this thesis use pictures or words as USs and pictures of forms, Kanjis, or objects as CSs.

Simultaneous versus Sequential Presentation. During the acquisition phase, CS and US can be presented simultaneously or sequentially. When presented simultaneously, CS and US appear at the same time and disappear together after a predefined time interval has passed. When presented sequentially, a temporal overlap of CS and US is possible (e.g., delay conditioning), but not necessarily given (e.g., forward conditioning). In this thesis, three types of conditioning were used: simultaneous conditioning, forward conditioning, and backward conditioning (see Figure 2.1).

Figure 2.1

Schematic Illustration of Simultaneous and Sequential Conditioning Procedures.



Note. The arrow illustrates the direction of time. “On” represents the stimulus onset, “Off” the stimulus offset. The trial course for the US is highlighted in grey. In simultaneous conditioning procedures, CS and US on- and offset are congruent. In forward conditioning procedures, US onset starts only after CS offset. Similarly, in backward conditioning procedures, the CS onset starts only after the US offset. Adapted from Mazur & Odum (2023).

Contiguity. The stimuli used in EC experiments are usually presented in close spatio-temporal proximity. Although research has shown that statistical contingency is important for classical conditioning (e.g., Rescorla, 1968), CS-US contingency seems to have less impact on EC. By manipulating the degree of contingency between CS and US, Baeyens and colleagues (1993) used four conditions to test this assumption: a perfect contingency condition with either ten or 20 CS-US pairings, a partial reinforcement condition, and a composite condition. In the perfect contingency condition, they implemented perfect contingency by presenting CS and US together ten (or 20) times. In the partial reinforcement condition, in addition to the ten CS-US presentations, they presented ten CSs alone. In the composite condition, ten CS-only and ten US-only were presented in addition to the ten CS-US pairings (i.e., in this last condition, contingency was absent). As the authors did not find a significant difference between conditions, they concluded that EC is insensitive to contingency: The CS does not have predictive value for the US (Baeyens et al, 1993; see Hofmann et al., 2010, for a meta-analysis). Other research has shown that EC effects can be found with all three procedures, irrespective of duration of stimulus presentation and length of interstimulus interval (e.g., Gast et al., 2016; Hütter & Sweldens, 2013; Kim et al., 2016).

Consequently, researchers proposed to discard contingency as a prerequisite for EC, and rely on contiguity instead (e.g., Martin & Levey, 1994). In the previous EC literature, contiguity is mostly defined as the spatial and temporal proximity of CS and US (e.g., De Houwer et al., 2020; Jones et al., 2010; Walther, 2002). Further evidence for the significance of contiguity comes for example from Jones and colleagues (2009), who observed larger EC effects when CS and US appeared in close spatial proximity. Additionally, it seems to be an unwritten law that CS and US do appear in close temporal proximity (i.e., most EC experiments show CS and US within an interval of 1500 ms, Gast et al., 2016). Thus, it is justified that environmental factors like spatial and temporal proximity are considered to be quite influential factors regarding EC and learning per se (cf. Chapter 1). While all experiments reported in this thesis present CS and US in spatio-temporal contiguity, implications of this proceeding are further discussed in Chapter 7.

2.2 The Organism

Yet, these environmental aspects do not solely determine the occurrence and size of an EC effect; factors on the side of the organism can also influence EC. The organism's importance for EC is already emphasized by an extensive number of EC studies (e.g., studies on contingency awareness, attention, processing goals, and memory; see, e.g., Gast & Rothermund, 2011a; Hütter et al., 2012; Stahl & Corneille, 2019). However, most of this previous research is guided by the question of which processes underlie EC (i.e., the *how* question). The present thesis, however, is concerned with the *when* question of EC. Yet, to better understand the organism's role for the *when* question, I will briefly address the processes that might underlie EC effects (i.e., the *how* question).

2.2.1 Process Theories

The question of *how* EC occurs has been the subject of debate (e.g., Hofmann et al., 2010; Hütter et al., 2012; Jones et al., 2010) and it remains unresolved (Moran et al., 2023). An important part of this question is the role of contingency awareness. Contingency awareness refers to "the phenomenon of becoming consciously aware of the association between a conditioned stimulus (CS) and an unconditioned stimulus (US)" (Baeuchl et al., 2019, p. 811), and is often measured by assessing contingency memory (i.e., reporting the knowledge of CS-US contingencies after the acquisition phase; Bar-Anan et al., 2010). As contingency memory usually captures knowledge about the contiguity of CS-US pairings (instead of knowledge about their statistical contingencies), Gast et al. (2012) proposed to use the term contiguity memory instead. Yet, independent of the term used to describe the phenomenon, different process theories make different assumptions regarding contingency awareness. Based on these assumptions, researchers often split process theories into 'associative' versus 'propositional' accounts (e.g., Hütter & Fiedler, 2016; for reviews see Hofmann et al., 2010; Jones et al., 2010). Associative accounts do not rely on contingency awareness, whereas propositional accounts do. 'Associative-Propositional' models, on the other hand, assume both associative and propositional processes to play a role in learning. Although memory models can be considered as part of propositional models (cf. Hütter, 2022), I will explain them separately. Accordingly, I categorize the process theories into four groups: (a) single-process associative theories, (b) single-process propositional theories, (c) dual-process theories, and (d) memory models.

The conceptual categorization account (e.g., Davey, 1994) stands out as an exception since it does not presuppose the formation of associations, but at the same time, it refrains from making explicit assumptions about propositional or memory traces. Instead, the account assumes that the co-occurrence of CS and US increases the salience of features that CS and US share, leading to a change in the categorization of the CS as either liked or disliked (Hofmann et al., 2010).

Associative Theories. Associative processes involve "the automatic formation of mental associations between co-occurring stimuli" (Gawronski & Bodenhausen, 2018, p. 2). Automaticity encompasses various aspects (see Moors & De Houwer, 2006, for a conceptual analysis); however, the 'for horsemen of automaticity', as Bargh (1994) labeled the most frequently recurring features in the capacity literature, are unconsciousness, unintentionality, uncontrollability, and efficiency (Moors & De Houwer, 2006). Thus, processes related to the formation of associations (in contrast to proposition formation) are believed to function without participants' awareness, without their intentional learning, beyond their control, and are processed rather quickly (De Houwer, 2018). Gawronski and Bodenhausen (2009; 2011) referred to these as the *operating conditions* (i.e, "the circumstances under which a certain process is assumed to operate", Hütter, 2022, p. 642; see also Corneille & Stahl, 2018), that they distinguish from the *operating principles*, which describe how the process works. There are currently three different accounts that are mainly discussed as single-associative process explanations for EC: the holistic account (e.g., Levey & Martin, 1987), the referential account (e.g., Baeyens et al., 1992), and the implicit misattribution account (Jones et al., 2009).

In the referential account, the associations are thought to be formed between CS and US (i.e., S-S learning). However, different from classical conditioning, the CS does not *predict* the US (i.e., signal learning), but *refers* to the US (i.e., referential learning, Baeyens et al., 1992). A rather similar assumption is made by the holistic account. This account proposes a holistic representation that contains features of CS and US as well as the US valence. The CS can associatively activate this representation and thus, the associated US evaluation (i.e., S-S and S-R learning might take place, Martin & Levey, 1994). The implicit misattribution account, on the other hand, postulates that source confusion is responsible for EC: the evaluative reaction evoked by the US is incorrectly attributed to the CS (Jones et al., 2009). Implicit misattribution can thus be considered a theory of S-R learning.

Propositional Theories. In propositional theories, such automatic processes of association formation between S-S or S-R are not assumed. Rather, these theories propose that participants must form conscious (non-automatic) propositions about CS-US relations. Propositions convey information about events in the world, and as such, can vary in their (perceived) degree of accuracy (De Houwer, 2009; Mitchell et al., 2009). Thus, propositions can be perceived as true, as false, or as something in between (i.e., they have a so-called truth value); even if their actuality cannot be determined (e.g., 'God is a woman'; cf. De Houwer, 2018; see also De Houwer et al., 2020). Propositional accounts of EC assume that participants form propositions during the acquisition phase and draw on these propositions in the evaluation phase (e.g., 'CS and US co-occur', 'Similar things tend to co-occur', 'The CS probably has the same valence as the US'; cf. Van Dessel et al., 2019; see also De Houwer, 2018). Additionally, some of these models make assumptions about the storage of propositions in memory (e.g., the Integrated Propositional Model, IPM; De Houwer, 2018).

Dual-Process Theories. Dual-process accounts like the APE model (Gawronski & Bodenhausen, 2006) combine associative and propositional learning mechanisms to explain EC (e.g., Gawronski & Bodenhausen, 2018; for a review see Hütter, 2022). Associations are mental representations that do not contain information about the relation between stimuli (i.e., they do not assess a truth value). Propositions, on the other hand, are mental representations with truth value that do contain such relational information (Hütter, 2022). Mental associations are thought to be driven by spatio-temporal proximity of stimuli and similarity-based retrieval (i.e., retrieval is based on feature similarity between input stimuli and accessible memory representations). Associative processes are assumed to occur first, and propositional processes then operate on these initially formed associations (Gawronski & Bodenhausen, 2011). For example, the co-occurrence of a CS and a US in close spatio-temporal proximity might lead to the formation of an association between these stimuli and the respective judgment of the CS in the direction of the US. Yet, if participants receive instructions indicating that the CS holds the opposite valence of the US it is paired with, the association might be declined due to propositional processes.

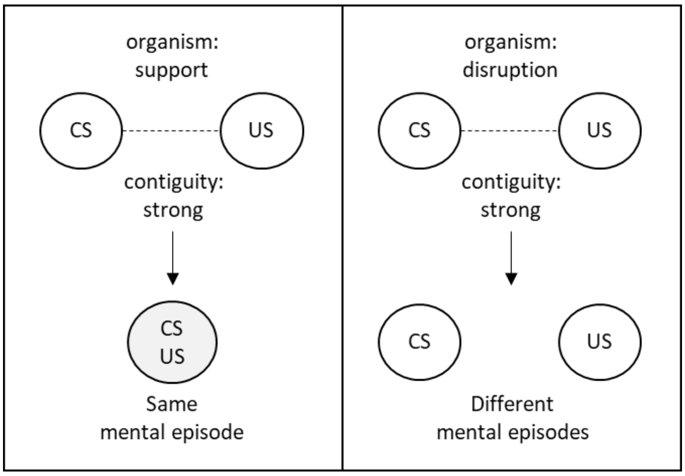
Memory Models. Memory-based models postulate EC's dependence on factors influencing the encoding, retention, and retrieval of episodic memory (Hütter, 2022; Gast, 2018; Stahl & Aust, 2018; Stahl & Heycke, 2016). The Declarative Memory Model (Gast, 2018) for example assumes that EC effects occur, if four conditions are met: (a) during the perception of CS-US pairings, memory traces are formed, (b) (part of) the trace links the CS to the US's evaluative information and is maintained until the CS valence is measured, (c) the (evaluative part of the) CS is retrieved and consciously represented during CS evaluation, and (d) the retrieved evaluative information is used for evaluating the CS.

2.3 The When Question

The above-listed theories merely dealt with the question of *how* EC occurs. As this thesis aims to answer the question of *when* EC occurs, I propose to differentiate between *co-occurrences* and *pairings*. A co-occurrence is typically, but not exclusively, the temporal and spatial contiguity of two stimuli; that is, their joint occurrence in the environment. However, for such a co-occurrence to become a pairing, the information from the environment requires the organism to perceive this co-occurrence. Figure 2.3 shows a simplified illustration of the organism's input (a more detailed illustration will be introduced in Chapter 4).

Figure 2.3

Schematic Illustration of the Organism's Input on the Formation of Mental Episodes.



Note. The dotted line represents a possible conjunction between CS and US (i.e., the possibility of encoding both stimuli into the same mental episode). The organism can support (left panel) or disrupt (right panel) this conjunction. Spatio-temporal contiguity of stimuli does not differ between panels. A supporting organism leads to the encoding of CS and US in the same mental episode. A disrupting organism prevents the formation of a shared mental episode. Each circle represents one mental representation (or mental episode).

As depicted, spatio-temporal contiguity of two stimuli (i.e., CS and US) can lead to a conjunction of the stimuli (left panel), but it does not necessarily have to (right panel). Thus, even though the stimuli co-occur (i.e., spatio-temporal contiguity is given), they might not be perceived as pairing (i.e., are not encoded into the same mental episode). To qualify as a pair, CS and US must be encoded into the same mental representation (or mental episode, cf. Chapters 4 & 5). It seems that most theories that try to explain which process inherently underlies EC assume the presence of mental representations. These mental representations are described as informational units stored in memory. However, even though these units are thought to be acquired through encoding processes (Hütter, 2022), none of these theories seems to ask *when* encoding takes place. Instead, some researchers use the term 'mental representation' synonymously with proposition (e.g., De Houwer, 2018). However, in this thesis, I define mental representations as simple mental constructs that can be enriched with any kind of information. Which type of information is represented in mental representations might depend on various factors. For example, one could assume that CS and US form a joint representation at a very early stage, which is amplified by a proposition (or an association) at a later point in time (for a similar argument regarding the storage of episodes in memory, see Stahl & Aust, 2018). Mental representation and mental episode will be used synonymously throughout this thesis.

To provide a concrete example, one may consider repeated walks in a dog park. There, observing a person repeatedly co-occurring with a cute dog should lead to an EC effect – the person should become more likable. Similarly, the environmental setup in the park implies the perception of dog and person as a pair; the observer should encode them into the same mental episode. However, if there is a leash policy in the park, but the person does not have a dog leash, the observer may perceive person and dog as not belonging together and may not encode them into the same mental episode.

Still, the encoding of two stimuli into the same mental episode allows for the absence of EC effects (e.g., when retrieval fails, cf. Stahl & Aust, 2018), and different processes to take place between the two stimuli. In the given example, the perception of dog and owner as a pair might lead to automatic associations between the stimuli, or the proposition may follow from the person's facial expression that the person likes or hates the dog.

Thus, the *when* question is not to be confused with the *how* question: The *when* question is concerned with the encoding of stimuli into mental episodes; the *how* question is concerned with the type of relation that these mental episodes may or may not contain (i.e., associations, propositional information, or both). The difference between *when* and *how* question will be discussed further in Chapter 4.

2.3.1 The Dual-Force Perspective

To answer the *when* question of EC, Sperlich & Unkelbach (2022) developed a dual-force perspective of EC (cf. Chapter 4). This perspective relies on the concept of mental representations. Whether CS and US are encoded into the same mental representation (i.e., whether CS and US are perceived as pairing), depends on the interplay of organism and environment. The assumption that these two forces are not independent, but interact with one another, is based on Fiedler's dual-force model (2001; Fiedler & Bless, 2000). Using Piaget's (1954) terminology, Fiedler defines *accommodation* as the adjustment of the cognitive system to the stimulus environment, and *assimilation* as the transformation of external information to fit into internal knowledge structures. Accommodation can easily be translated to bottom-up (i.e., stimulus-driven) effects that influence the organism while assimilation can be construed as top-down (i.e., organism-driven) effects.

In typical EC experiments, CS and US are presented in close spatio-temporal proximity (i.e., spatio-temporal contiguity). However, even though contiguity is allocated at the bottom-up (i.e., the environmental) side of the dual-force model, it is based on a principle known from Gestalt psychology: the law of proximity. The proximity of two stimuli was described by Wertheimer (1938) in terms of their auditory organization: "tap-tap, pause, tap-tap, pause" (p. 74). In EC experiments, this organization is mirrored: CS-US, blank, CS-US, blank. As following such rhythmic patterns is almost inevitable for the organism (Neisser, 2014), it likely supports the bottom-up structure of CS-US presentation, thereby forcing CS and US into the same mental episode. However, the organism might not always support such bottom-up patterns. Potentially, there are even situations in which the organism prevents encoding CS and US into the same mental episode. A test for this assumption will be provided in Chapter 4. Additionally, Chapter 5 provides further evidence for the significance of top-down input under conditions of consistent spatio-temporal contiguity.

The present distinction in its explicit form was absent from most EC research, which answered the *when* question from a behavioristic perspective, defining the occurrence of CS and US close together in space and time as the sufficient antecedent (e.g., Hütter et al., 2012; Jones et al., 2010). However, the same appearance of two stimuli may or may not lead to people's construal of the stimuli as a pair, depending on features of the environment or internal processes applied to the stimuli. Based on its roots in learning psychology, EC research addressed co-occurrences either from a learning perspective, for example, by investigating CS-US presentation schedules (e.g., Kattner et al., 2012) or reinforcement schedules (e.g., Kattner, 2014). If research addressed the organism, the focus was on cognitive processes during or after the co-occurrence, for example, the role of memory or attention as predictors for EC effects (e.g., Corneille & Stahl, 2019; Stahl et al., 2009). From a social-cognitive perspective, the distinction between bottom-up input (e.g., spatial and temporal contiguity) on the environmental level and top-down input (e.g., attention) on the organismic level seems less surprising. The insight that the psychological reality (here, the organism's input) and not the physical reality (here, the environmental input) determines behavior was already famously formulated by Lewin (1936). There is already evidence for the relevance of this organism-environment interaction in EC. For example, Gast and De Houwer (2012) showed that one might inform participants that a stimulus will appear together with something positive or something negative: Participants liked the stimulus that they expected to be followed by something positive more compared to the stimulus they expected to be followed by something negative. However, this has been considered mainly as evidence for EC's underlying processes, but not for the question of *when* EC should occur.

Chapter 3: The Evaluative Information Ecology

The focus of the previous chapter was on the experimental structures of EC. These structures can be manipulated to some extent. For example, the number of stimuli can be varied, the proximity of one stimulus to the other can be adjusted, one type of stimuli can be replaced with another, and the context in which the stimuli occur can be changed. This artificial environment of an experiment does, however, not necessarily portray reality. Even though EC experiments are regarded as a laboratory model for the development of attitudes in many real-life situations (Hütter & Tigges, 2019), the real world is more complex than the experimental environment (Hütter et al., 2014). Additionally, despite keeping components that are not part of the hypothesis constant, factors within the organism may still vary. For example, the organisms might be pre-disposed to learn some combinations of CS and US more readily than others (Fiedler & Unkelbach, 2011). This *preparedness* notion (cf. Seligman, 1970) is based on phylogenetic assumptions (cf. Chapter 1), which are also implicated in explaining the advantages of negative over positive information (i.e., the so-called negativity bias). Such an *evolutionary approach* suggests that the stronger impact of negative information compared to positive information is caused by evolutionary circumstances. Conversely, Unkelbach and colleagues (2019) explain occurrences of negativity bias based on an ontogenetic perspective using a *cognitive-ecological approach.* This approach emphasizes the influence of the environment (i.e., bottom-up input) on biases. In the next chapter, I will describe both approaches in more detail, review relevant literature, and explain why there is no negativity bias in EC experiments.

3.1 Valence Asymmetries

Even though most people distinguish easily between 'good' and 'bad' information, these two classes are not produced equally, but differ in the way they are processed. Such valence asymmetries (e.g., Kanouse & Hanson, 1972; Peeters, 1971; Peeters & Czapinski, 1990) are expressed in an advantage of negative over positive information (i.e., a *negativity bias*), or an advantage of positive over negative information (i.e., a *positivity bias*[6] or *pollyanna effect*; e.g., Boucher & Osgood, 1969).

[6] But see Unkelbach et al., (2020), for a critical discussion of the terms 'advantage' and 'bias'.

Importantly, this advantage occurs even though the number of positive and negative information is held constant.

In the following section, I will first explain these valence asymmetries in more detail. I will then address the difference between affective and ecological explanations for the negativity bias with special attention to the Evaluative Information Ecology (EvIE) model put forward by Unkelbach and colleagues (2019). Finally, I will transfer the gained insights to EC experiments.

3.1.1 Positivity Bias

As mentioned before, a positivity bias describes the advantage of positive information over negative information. For example, the classification of attitude objects as good (Bargh et al., 1992) and the evaluation of positive words (Klauer & Musch, 1999; Unkelbach et al., 2008) were faster than the classification or evaluation of their negative counterparts, and nonsense syllable words were learned better when paired with a pleasant word (vs. an unpleasant word; Anisfeld & Lambert, 1966). Using the Sorting Paired Feature (SPF) Task, associations with positive valence were more strongly related to self-report than associations with negative valence (e.g., Bar-Anan et al., 2009). Additionally, mean scores for Positive Affect were higher than mean scores for Negative Affect (Brans et al., 2013; for reviews see Unkelbach et al., 2019; 2020). Finally, in a more recent study, similar positive attitudes tended to generate more liking than comparable negative attitudes (Zorn et al., 2022). However, Unkelbach and colleagues (2020) refer to positivity bias as a step-child regarding valence asymmetries. Baumeister and colleagues (2001) make a similar statement when it comes to the higher publication rate of research articles that cover negative issues: “More positive phenomena had to wait until the recent emergence of stronger methods, more sensitive measures, and better statistical techniques” (p. 324). In contrast, the notion that ‘bad’ is stronger than ‘good’ is evident in various areas of psychology (Baumeister et al., 2001).

3.1.2 Negativity Bias

As previously mentioned, this general tendency of individuals to give stronger weight to negative information is called negativity bias (Rozin & Royzman, 2001). Losing money leads to greater distress than the joy gained by winning money as proposed by the prospect theory (e.g., Kahnemann & Tversky, 1984). Additionally, negative interactions are stronger predictors for marital satisfaction than positive interactions (e.g., Gottman & Krokoff, 1989), people recall negative events more often than positive events (e.g., Finkenauer & Rimé, 1998), angry faces are detected more rapidly than happy faces (e.g., Hansen & Hansen, 1988), and negative stereotypes are formed more easily than positive ones (e.g., Rothbart & Park, 1986), to name just a few examples (for reviews see Baumeister et al., 2001; Rozin & Royzman, 2001; Unkelbach et al., 2020). Of special interest for this thesis, however, are the findings regarding negativity biases in learning research. Thus, I will review some of the literature on negativity bias in learning research; specifically, in operant conditioning, classical conditioning, and EC.

Operant Conditioning. In paradigms using operant conditioning, for example, it has been shown that negative information carries more weight than positive information: Loud tones, verbal punishments, and losing marbles elicited stronger effects than candy, verbal rewards, and gaining marbles (e.g., Constantini & Hoving, 1973; Meyer & Offenbach, 1962; Penny & Lupton, 1961). In a more recent study, participants' task was to unscramble anagrams; some of them solvable (e.g., "SEUMO" to "MOUSE"), others considered to be unsolvable due to their high difficulty (e.g., "UDARIVMIQU" to "QUADRIVIUM"). In the positive frame (i.e., reward) condition, they were told to gain $0.25 for every correct answer (with a maximum of $1.50 to 'win'); in the negative frame (i.e., punishment) condition, they started with $1.50 and lost $0.25 for every anagram they were unable to solve. Results showed that participants in the negative frame condition spent more time trying to solve the 'unsolvable' anagrams. Thus, the stronger effect of negative information (i.e., of punishment) was evident in the longer persistence of participants in the respective condition (Goldsmith & Dhar, 2013). The authors argue that the effect occurred due to a negativity bias in information processing. Similar results were found with approach-avoidance behavior as the dependent variable (e.g., Atthowe, 1960; Myers et al., 1961; Fazio et al., 2004).

Classical Conditioning. Classical conditioning procedures were used to examine learning with biologically significant stimuli, also known as preparedness (Seligman; 1970; Seligman & Hager, 1972). For example, in a study by Öhman, Erixon et al. (1975), participants received electric shocks which were either administered together with the presentation of neutral pictures (e.g., houses) or with biologically significant stimuli (i.e., snakes). The difference in participants' skin-conductance response (i.e., sweating) between neutral and phobic pictures was evident before shocking and persisted during conditioning: Pre-UCS as well as CS responses were stronger for snake pictures. Their main dependent variable, however, was extinction; that is, participants' skin conductance towards the pictures without electric shocks after the conditioning trials. Within five exposures to the house pictures after learning, participants' responses returned close to baseline. For the snake pictures, however, the skin conductance response proved to be relatively resistant to extinction (see also Hugdahl & Kärker, 1981; Öhman, Eriksson et al. (1975). Based on these and similar findings, Öhman and Mineka (2001) proposed a special *fear module*.

In addition to fear, disgust appears to be a very robust trigger of negative experiences (cf. Rozin & Royzman, 2001). Just as fear, disgust stimuli might be biologically relevant, as disgust protects us from infectious diseases (Curtis et al., 2004). As Rozin and Royzman (2001) already noticed: "Brief contact with a cockroach will usually render a delicious meal inedible. The inverse phenomenon – rendering a pile of cockroaches on a platter edible by contact with one's favorite food – is unheard of" (p. 296). Rozin and colleagues (1986) showed that dropping a sterilized dead (i.e., objectively not contagious) cockroach into a glass of juice decreased liking ratings of this juice, and even liking ratings of a new glass containing the same kind of juice. Importantly, that new glass of juice has never been in contact with the cockroach. The contact of juice with a candleholder, however, did not affect the ratings in a similar manner. The authors called their procedure *roach conditioning*, in reference to classical conditioning.

Evaluative Conditioning. As mentioned before, some researchers argue that EC is merely a subform of classical conditioning (e.g., Davey, 1994; Hofmann et al., 2010). There is, however, evidence that EC is qualitatively distinctive from classical conditioning (e.g., Bayens et al., 1992;

Baeyens & De Houwer, 1995; Kattner & Green, 2015). For example, EC as an attitudinal process does not refer to the prediction of events, but “to the affective or cognitive meaning attitudinal objects acquire in the context of pleasant or unpleasant experiences” (Walther et al., 2005; p. 176). Such pleasant or unpleasant experiences are not, as in classical conditioning, restricted to biologically relevant UCSs (e.g., food/electric shocks, Hütter, 2022). Instead, EC paradigms usually use USs such as pictures (e.g., Levey & Martin, 1975), words (e.g., Zanon et al., 2014) or sounds (e.g., Moran & Bar-Anan, 2013). Yet, there is some evidence for negativity bias in EC. For example, in their first paper on EC (cf. Chapter 2), Levey and Martin (1975) note:

> The results of interest to the conditioning hypothesis are the ratings of the first stimulus of the NEUTRAL-LIKED pair and the NEUTRAL-DISLIKED pair. The effect of negative evaluation was clearly stronger than that for positive evaluation, and this is consistent with our knowledge of aversive conditioning. (p. 224)

Similarly, Baeyens and colleagues (1990) found stronger effects for conditioning procedures using an unpleasant vs. a pleasant-tasting substance. However, the re-analysis of meta-analytical data shows no evidence of negativity bias (see Chapter 6 for more details).

3.1.3 Explanations of Negativity Bias

In previous research, negativity bias was often explained through the influence of valence-induced affect. That is, differences in the processing of negative and positive information were thought to depend on the energetic potential of an information (e.g., due to its relevance for survival; Alves et al., 2015).

Phylogenetic Explanations. This view is also held by Baumeister and colleagues (2001), who particularly emphasize the adaptive character of negative stimuli. From their perspective, better adaption to bad things leads to a survival advantage and thus, enables the organism to pass on its genes. In line with this notion, Rozin and Royzman (2001) note that death terminates the opportunities for reproduction and thus, must be prioritized in the evolutionary scheme. Put differently, if you miss the sabretooth tiger in the bushes, you will not pass your genes into the next generation; if you miss a mating opportunity, there will be many other possibilities. Consequently, negative information might grab attention (e.g., Pratto & John, 1991), sticks out in the field of vision (e.g., Hansen & Hansen, 1988; Öhmann et al., 2001), or has a lower perceptual threshold (e.g., Nasrallah et al., 2009). Thus, negativity bias as an evolutionary byproduct is explained through intra-psychic phenomena.

Ontogenetic Explanations. Unkelbach and colleagues (2020), however, point to the fact that phylogenetic explanations fail to explain apparent positivity advantages. Additionally, attempting to explain both, negativity and positivity bias might lead to predictions becoming rather arbitrary. The example just mentioned can serve as an illustration of this arbitrariness: Evolutionary pressure may be interpreted to work for negativity advantages for one case in one context (e.g., higher sensitivity for predators in the context of a sufficient number of mating partners), but for positivity advantages in another context (e.g., higher sensitivity for mating partners in a context with too little mating partners). Thus, other accounts are introduced to explain differences in the processing of positive and negative information (for a review see Unkelbach et al., 2020). In contrast to evolutionary explanations that follow a phylogenetic rationale, these explanations have an ontogenetic basis (i.e., the organism's behavior changes due to the environment's regularities, De Houwer et al., 2013). Two of these explanations (i.e., frequency and similarity) are addressed by the Evaluative Information Ecology (EvIE) model (Unkelbach et al., 2019). As the frequency approach is particularly important for this thesis (Chapter 6), I will use the following section to explain the EvIE model and related findings in more detail.

3.2 The Evaluative Information Ecology (EvIE) Model

The EvIE model makes two assumptions about the evaluative information in people's environment: (a) positive information is more prevalent than negative information (e.g., Alves et al., 2018b; Matlin & Stang, 1978; Unkelbach et al., 2010), and (b) positive information is more similar to other positive information than negative information is to other negative information (e.g., Alves et al., 2016; Koch et al., 2016; Unkelbach et al., 2008). Thus, the focus of the EvIE model lies on the structural properties of evaluative information (i.e., positive, or negative information) in the social environment. Referring to the model as ecological aims to distinguish it from earlier approaches that relied on intra-psychic explanations (e.g., motivational, or affective components within the organism) to account for differences in frequency and similarity of evaluative information. As opposed to these intra-psychic explanations, the ecological approach assumes that the differences between positive and negative information are truly present in the people's social environment. Yet, it is crucial to note that neither the organism nor intra-psychic events are neglected in this ecological approach. However, instead of using other intra-psychic phenomena (e.g., more attention to negative information) to explain intra-psychic phenomena (e.g., more differentiation of negative information), the model explains these intra-psychic phenomena with "an extra-psychic property of the ecology" (Unkelbach et al., 2019, p. 220).

3.2.1 The Evaluative Information Ecology

Unkelbach and colleagues (2019) defined the information ecology in line with Brunswik (1955) as an "objective, external potential offered to the organism" (p.198). Information within the ecology is then defined as anything that might be evaluated as 'good' (i.e., positive) or 'bad' (i.e., negative) by the organism, and can be factual (e.g., experiences) or symbolic (e.g., language). Without the organism's assessment, observable conditions are deemed to be neither positive nor negative. Unkelbach and colleagues (2019) thus differentiate between *substance* of reality and *evaluation* by the organism (see also Leising et al., 2015).

The substance of reality might be something like the salinity of water, a beverage's degree of heat, or the amount of newly fallen snow. The inherent nature of these substances is neutral. Their characterization as good or bad only arises through the process of evaluation by the organism. For instance, salty water proves favorable for a herring, yet unfavorable for a trout; and vice versa. A cold beverage might be refreshing in warm weather, but undesirable when it's cold; the opposite holds true for a warm drink (Unkelbach et al., 2019). Similarly, plenty of snow might be hoped-for on a ski vacation but frustrating if it delays the departure to a vacation destination; conversely, the lack of snow is beneficial for a scheduled departure but not ideal for a day of skiing. The substance is thus considered the underlying factor for the structural properties on the level of evaluations. Put differently, differences in the frequency and similarity on the substance level should also be represented on the evaluation level (Unkelbach et al., 2019).

3.2.2 Evidence for the Evaluative Information Ecology Model

The EvIE model assumes that positive information is more frequent than negative information in the information ecology. To validate this assumption, Unkelbach and colleagues (2019) reviewed different studies. As language reflects reality to a certain degree (Unkelbach et al., 2019), they also included studies from the psycho-lexical domain and found that nouns are more often classified as positive than negative (Boucher & Osgood, 1969), positive words are used more often than negative words (Augustine et al., 2011), and number as well as occurrence is higher for positive than for negative words (Warriner & Kuperman, 2015). In addition to this higher frequency of positive information, the EvIE model assumes that positive information is more similar than negative information. The strongest evidence comes, according to Unkelbach and colleagues (2019), from emotion research. Most theories of basis emotions postulate more negative (e.g., anger) than positive emotions (e.g., joy) and thus point towards a greater diversity of negative emotions. For example, only two of the 12 presented sources on emotion theories propose an equal amount of positive and negative emotions (Mowrer, 1960; Weiner & Graham, 1984), whereas the remaining ten theories assume more negative than positive emotions (e.g., Ekman et al., 1982; James, 1884; Plutchik, 1980). Together with other studies discussed in their review, Unkelbach and colleagues (2019) provide a substantial body of evidence for their assertion of the higher prevalence and similarity of positive information in the ecology.

3.2.3 Explanations for the Higher Frequency of Positive Information

There are different assumptions about the causes for the higher frequency of positive information in the evaluative ecology. In their review, Unkelbach and colleagues (2019) discuss evolutionary necessities, reinforcement learning, and hedonic sampling.

Evolutionary Necessities. The idea that the higher frequency of positive information has evolved evolutionary is based on the principle that social life has survival advantages when organizing itself in a cooperative instead of a competitive way. Evidence for such a cooperative organization comes from economic games. Using the *prisoner's dilemma*[7], Axelrod and Hamilton (1981), for example, suggested that, in terms of fitness, defection is the best strategy if the game is only played once. However, individuals do not usually engage in just a single interaction; instead, they meet repeatedly. Yet, the exact number of encounters is, in most cases, not predetermined. The best strategy for this iterated prisoner's dilemma is then not defection, but cooperation. Given that most individuals perceive cooperative behavior as positive and defection as negative, Unkelbach and colleagues (2019) suggest that cooperation leads to positive experiences that should be prevalent at least on a societal level (i.e., single individuals might still show defecting behavior).

Reinforcement learning. In addition to the evolutionary attempt, Unkelbach and colleagues (2019) propose an approach based on Thorndike's (1898) *law of effect*. The authors assume that positive behavior is exhibited more frequently because it is rewarded, whereas negative behavior is shown less frequently because it is punished. Even though the evolutionary explanation and the reinforcement approach are not independent (i.e., why positive behavior is rewarded, and negative behavior is punished could be explained by making evolutionary assumptions), they are located at different explanatory levels: phylogenetic (evolutionary necessities) and ontogenetic (reinforcement learning; cf. De Houwer et al., 2013).

[7] In the prisoner's dilemma, two participants face the decision to either cooperate or defect. If both choose cooperation, they receive a reward (such as money). If both opt for defection, none of them is rewarded (e.g., they do not receive the money). If one player defects while the other cooperates, the defector obtains the reward (e.g., receives all the money), and the cooperating player gets nothing (Axelrod, 1980).

Hedonic Sampling. Finally, the prevalence of positive information might follow from hedonic sampling. That is, people are actively looking for positive information (e.g., friendly people, fun events, nice clothes) while avoiding negative experiences (Denrell, 2005; Fazio et al., 2004). One special feature of the hedonic sampling suggestion is that it does not require the additional assumption of an information asymmetry on substance level (i.e., objectively more positive than negative information). Even in an ecology in which positive and negative information is equally frequent (an assumption that the EvIE does not make), the evaluative environment should be positive, since sampling is not random but based on hedonistic principles (Unkelbach et al., 2019).

3.2.4 Explanations for the Higher Similarity of Positive Information

To explain the assumed higher similarity of positive information, Unkelbach and colleagues (2019) discuss three further explanations: co-occurrences, the range principle, and affective influences.

Co-Occurrences. It seems evident that information that occurs more frequently must also *co-occur* more frequently. Thus, based on the assumption that positive information occurs more frequently than negative information, Unkelbach and colleagues (2019) computed the probability of co-occurring positive versus co-occurring negative hypothetical information. For example, if 70% of all individuals fall within the *good* range regarding emotional stability and extraversion, and 15% have *too much* or *too little* of each, then the probability of a person being in the 'good' range for both personality traits is p = .70 x .70 = .49. Conversely, the probability of a person being in the 'bad' range for both traits is p = .15 x .15 = .0225. Thus, the probability of positive information to co-occur is about 20 times higher than the probability of negative information to co-occur (see also Alves et al., 2016). As the higher frequency of co-occurrences is related to subjective interstimulus similarity (e.g., Griffiths et al., 2007) and stronger associations in memory (e.g., Fiedler et al., 2013), Unkelbach et al. (2019) infer that positive information should be perceived as more similar than negative information. In other words, the higher frequency of co-occurring positive information leads to higher subjective similarity of positive information.

Range Principle. However, similarity perception might also be explained in a more objective manner. While the argument for *subjective* similarity is based on the assumption that positive information is more frequent than negative information, the consideration of *objective* similarity bears on the "well-documented assumption that valence is a function of attribute extremity" (Alves et al., 2017b, p. 72). That is, when mapping different attributes on a range, positive attributes are usually located in the middle of this range, whereas negative attributes mark the extremes. For example, a Celsius thermometer typically ranges from -30 to +50 degrees. Thus, comfortable (i.e., positive) temperatures are located toward the middle of the range, whereas *too cold* and *too hot* (i.e., negative) temperatures can be found toward the ends of the range. Consequently, the maximum distance between two positive attributes must be smaller than the maximum distance between two negative attributes (Alves et al., 2017b). Translated to social attributes, a person can be *too little* (defect) or *too much* (excess); two character traits that frame a desirable personality (i.e., the mean). For example, to be perceived as courageous in a positive way, two people must lie within the medium range; conversely, courage as a negative attribute can be displayed in two rather distinct ways: someone who is *too timid* (i.e., has too little courage) becomes a coward (defect), and someone who is *too adventurous* (i.e., has too much courage) becomes reckless (excess, Aristotle, 1999; see also Alves et al., 2017b). Thus, as the "middle range typically meets the needs of an organism, […] the higher similarity within this 'positive' range follows" (Unkelbach et al., 2019, p. 240).

Affective Influences. A third explanation that Unkelbach and colleagues (2019) put forward is affective influences. Previous research showed that (phasic) affective influences (i.e., affective states only lasting for a few seconds) modulate the breadth of spreading activation from a prime to a related target. That is, priming was facilitated for prime-target pairs (e.g., KING – CROWN) when accompanied by a positive stimulus (e.g., a happy face) and inhibited when accompanied by a negative stimulus (e.g., a sad face, Topolinski & Deutsch, 2013). Contrary to the EvIE model, Topolinski and Deutsch (2013) argue that, in the information ecology, positive and negative information are equally similar. Differences in similarity then occur due to affect-induced processing differences.

Positive stimuli elicit shallow processing, and negative stimuli trigger deep processing, which in turn leads to high versus low perceived similarity (Alves et al., 2017a). However, when directly comparing this intra-psychic explanation with an ecological explanation (i.e., higher density of positive stimuli in the ecology; see also Unkelbach et al., 2008), no such influence of affect on perceived similarity was found (Alves et al., 2018a; see also Unkelbach et al., 2019).

3.2.5 Artificial Ecologies

Following the argumentation of phylogenetic and ontogenetic explanations for valence asymmetries, one might conclude that all experiments containing positive and negative information should demonstrate positivity biases, negativity biases, or both. However, when re-analyzing meta-analytic data of EC experiments, no evidence for any kind of bias was found (Chapter 6). As mentioned before, the EvIE model regards reality as a substance, that just exists (i.e., is neither good nor bad without evaluation). Yet, this substance consists of more information that will be evaluated positively (vs. negatively) by most people. Thus, the structural properties of the substance determine the structural properties of the evaluation (Unkelbach et al., 2019). Put differently, the EvIE model assumes that the evaluative information ecology and the social environment are equal regarding structural features of evaluative information. However, due to their symmetrical construction, this assumption might not hold true for artificial evaluative ecologies such as experiments.

Similarity in EC. The information ecology in EC experiments represents such a highly artificial construction. Thus, the similarity assumption mentioned earlier, might not be valid for EC experiments, as they differ from the EvIE model's assumed ecology in at least two important aspects: (a) positive information usually does not co-occur more often with other positive information, and (b) selection of positive and negative USs is typically based on preratings (i.e., researcher try to equate the affective potential and extremity for USs across categories). Put differently, researchers artificially adjust the positivity and negativity of USs, reducing the 'genuine' advantages in the organism's processing of negative USs. Thus, stimuli used in EC experiments might not differ in their similarity, as they were inadvertently selected to be comparable. Therefore, in the artificial ecology of an EC experiment, the assumption that positive information is more similar than negative information, might not apply. In this thesis, however, I have focused solely on examining frequency in EC experiments.

Frequency in EC. Many EC studies use pictures as USs, for example from the International Affective Pictures System (IAPS; Lang et al., 2008; e.g., Alves & Imhoff, 2023; Houben et al., 2010; Pleyers et al., 2009). Based on findings from Koch and colleagues (2016), IAPS pictures should lead to valence asymmetries in EC. However, EC experiments differ in yet another significant aspect from the evaluative ecology of the EvIE model: In a typical EC experiment, a determined number of CSs is presented *equally* often with positive and negative USs. Thus, in contrast to the asymmetrical evaluative ecology suggested by the EvIE model, the ecology's structure in EC experiments is symmetrical. Such a symmetric setup only allows for a frequency-induced negativity bias if additional assumptions are included. For example, if participants transfer the environmentally learned asymmetric information distribution to the constructed ecology of the EC experiment (cf. Unkelbach et al., 2019). Without that assumption, the EvIE model does not predict a negativity bias in EC. Yet, the EvIE model would predict valence asymmetries for an asymmetric EC setup (see Figure 3.2).

Figure 3.2

Schematic Illustration of EC Experiments with Different Frequencies of Positive and Negative Information.

pos pos pos pos **neg** pos	neg neg neg neg **pos** neg

Note. The left panel illustrates an EC experiment where positive information (pos) is frequent and negative information (neg) is rare. The right panel depicts the opposite structure. Adapted from Unkelbach et al. (2020).

If positive information is frequent and negative information is rare, the advantage of negative information should be expressed in the more negative evaluation of negative information (as compared to when negative information is frequent). Similarly, positive information might be evaluated more positively when infrequent (as compared to frequent). A test of this assumption is provided in Chapter 6.

3.3 Summary and Predictions

Learning can be defined in a phylogenetic way as the adaptation to the environment over generations, or in an ontogenetic way as the adaptation to environmental regularities during an organism's lifespan (De Houwer et al., 2013). One type of learning based on such an ontogenetic definition is EC. Three types of evaluative learning can be distinguished based on such an ontogenetic definition: "(a) regularities in the presence of one stimulus, (b) regularities in the presence of two stimuli, and (c) regularities in the presence of behavior and stimuli" (De Houwer & Hughes, 2020, p. 2). EC belongs to the second type of learning and can be defined as the change in the valence of a CS due to its pairing with a valenced US (De Houwer, 2007). In this thesis, I propose to distinguish between a *pairing* and the *co-occurrence* of stimuli. The co-occurrence of stimuli relies on their spatio-temporal contiguity (i.e., stimulus-driven bottom-up input), whereas pairings involve top-down input (i.e., input from the organism). If bottom-up and top-down work together, CS and US are forced into the same mental episode (i.e., are perceived as a pairing). If, however, bottom-up input is weak (cf. Jones et al., 2009), or top-down input works against bottom-up input, CS and US should not be encoded into the same mental episode (cf. Chapter 4). Additionally, strengthening bottom-up and/or top-down input should facilitate the encoding of CS and US into the same mental episode. We tested this assumption in Chapter 5. In these rather typical EC experiments, positive and negative information is presented equally often. Based on an evolutionary approach, organisms should be pre-disposed to learn negative information more readily than positive information. Yet, such a top-down influence seems to be absent in EC experiments; a finding that is rather surprising considering the evidence for negativity biases in other domains (Baumeister et al., 2001; Rozin & Royzman, 2001; Unkelbach et al., 2019). However, taking the cognitive-ecological approach into account suggests that the symmetrical structure of EC experiments might contribute to the absence of negativity bias in EC. Thus, in Chapter 6 we investigate the impact of information manipulation based on predictions of the

EvIE model (Unkelbach et al., 2019): In EC experiments that mirror the evaluative information ecology of the social environment (i.e., positive information is more frequent than negative information), a negativity bias should occur. If, however, the information ecology is manipulated in a way that negative information is more frequent, a positivity bias should occur. Thus, by adopting a dual-force perspective and incorporating both, bottom-up and top-down input, I will try to answer two questions in the remainder of this thesis: (a) *when* do co-occurrences lead to likes and dislikes, and (b) why is there no negativity bias negativity in EC? In a broader sense, answering these questions might help to understand why EC effects are consistently found in the laboratory, but rarely occur in the real world. I will further discuss this observation, as well as our findings, their implications, and possible limitations in the General Discussion in Chapter 7.

Chapter 4: The Introduction of a Dual-force Perspective on EC

The conditions determined by the experimental environment are considered an important component of EC. This becomes especially evident in the definition of EC as the change in the valence of a CS due to its *pairing* with a positive or negative US (De Houwer, 2007). This pairing is often equated with the spatial and/or temporal contiguity of CS and US (e.g., De Houwer et al., 2001; Jones et al., 2010; Rydell & Jones, 2009). Thus, a pairing is confounded with the *co-occurrence* of two stimuli. Based on a dual-force assumption, the spatio-temporal contiguity of CS and US provides the stimulus-driven bottom-up input of an EC experiment. The complementary force is illustrated by the top-down-input (i.e., the organism's influence). In typical EC experiments, these two forces are of different strengths: The bottom-up input represents the stronger force but is supported by the (relatively weak) top-down input. If, however, the top-down input counteracts the bottom-up input, co-occurrences might not be equivalent to pairings. That is, under conditions of consistent spatio-temporal contiguity (i.e., co-occurrence is given), CS and US might not be encoded in the same mental episode (i.e., they are not perceived as a pairing). We tested this assumption in the following chapter.

This chapter is based on the following article:

Sperlich, L. M., & Unkelbach, C. (2022). When do people learn likes and dislikes from co-occurrences? A dual-force perspective on evaluative conditioning. Journal of Experimental Social Psychology, 103, 104377. https://doi.org/10.1016/j.jesp.2022.104377

Please note that certain modifications were made to the headings, citation style, and formatting to align with the layout of this dissertation. No changes were made to the content of the article.

When do People Learn Likes and Dislikes from Co-Occurrences? A Dual-Force Perspective on Evaluative Conditioning

Abstract

Evaluative Conditioning (EC) research shows that people learn their likes and dislikes due to the co-occurrence of stimuli (CS and US) in the environment. Most recent EC research addressed processes underlying this phenomenon: how do people acquire their likes and dislikes? We address the question of when people learn from co-occurrences. To understand when learning occurs rather than how it occurs, we apply a dual-force perspective (Bless & Fiedler, 2006). We propose that the environment provides a bottom-up force, for example, via spatial and temporal closeness of CS and US. The organism provides a top-down force, for example, by active inferences or existing knowledge about CS and US. We propose that these forces jointly determine when people learn from co-occurrences. We tested this dual-force perspective by creating experimental conditions that should lead to EC effects from both associative and propositional EC process models but should prevent EC effects from a dual-force perspective. Across four pre-registered laboratory experiments (N = 568) and two pre-registered online experiments (N = 440), we found that changing the top-down force weakened and eliminated (Experiments 1-3) or reinstated EC effects (Experiment 4), using a backward (Experiments 1-4) or forward conditioning procedure (Experiment 5-6), while keeping the environmental bottom-up force constant (Experiment 6). Independent of the how question (e.g., associative or propositional), these results support the dual-force perspective as a framework to predict when people learn their likes and dislikes from the co-occurrence of CS and US in the environment.

Keywords: Evaluative conditioning, Attitude acquisition, Evaluative learning, Environment-organism interactions, Dual-force model

4.1 Introduction

How people acquire their likes and dislikes is a prime research topic in social psychology. One explanation that has received substantial attention is Evaluative Conditioning (EC). A stimulus, called a conditioned stimulus (CS), becomes liked or disliked due to its co-occurrence with another likable or dislikeable stimulus, called an unconditioned stimulus (US). The interest in EC follows from the phenomenon's broad implications across theoretical approaches and domains, from basic experimental research to applied questions of consumer psychology. The CS may be a person (e.g., Fiedler & Unkelbach, 2011), a social group (e. g., Olson & Fazio, 2006), a consumer good (e.g., Walther & Grigoriadis, 2004), a brand (e.g., Sweldens et al., 2010), or a simple abstract stimulus (e.g., Högden et al., 2018). The US may be a picture (e.g., Alves et al., 2020), a word (e.g., Dijksterhuis, 2004), a sound (e.g., Moran & Bar-Anan, 2013), or a smell (e.g., Baeyens et al., 1996). Thus, the simple co-occurrence of two stimuli could explain the acquisition of attitudes towards people, groups, or consumer products.

EC research was facilitated by De Houwer's (2007) definition of EC as an *effect*, namely "an observed change in liking that is due to the pairing of stimuli" (p. 233). This definition established a criterion of what should be considered EC, and it allowed researchers to consider, test, and compare various underlying process models that explain observable changes in liking (see Bar-Anan & Balas, 2018). These process models included propositional reasoning (e.g., De Houwer, 2018; Mitchell et al., 2009), misattribution (Jones et al., 2009; March et al., 2018), memory (Gast, 2018; Stahl & Aust, 2018), and associative processes in combination with propositional reasoning processes (Gawronski & Bodenhausen, 2018; Hu et al., 2017a). These models make differential predictions for the role of memory for the pairings in EC (e.g., Gast et al., 2012; Mierop et al., 2017), the role of awareness of the pairings (e.g., Hütter et al., 2012; Stahl et al., 2009), and to some extent, for the outcomes on direct and indirect evaluative measures (e. g., Gawronski et al., 2015; Moran & Bar-Anan, 2013).

These and many more investigations thereby addressed *how* people learn likes and dislikes from the co-occurrence of stimuli (see De Houwer et al., 2020; Jones et al., 2010; for narrative summaries). Here, we approach the EC phenomenon from a different perspective and ask: *When* do people learn evaluations from co-occurrences of two stimuli? This question has garnered much less interest compared to the how question, although it is one of the earliest puzzles faced by EC researchers: if the mere co-occurrence of two stimuli suffices for EC effects, why are EC effects not observed more frequently in ecological settings (see Jones et al., 2010; “a seemingly ephemeral effect”, p. 209; Rozin et al., 1998)? In the following, we aim to answer this question.

To be sure, there is substantial research that added pre-conditions on the organism’s side to the mere co-occurrence of a CS and US in space and time. Most prominently, the question of whether people need to be aware of the co-occurrence (i.e., ‘contingency awareness’) stimulated much research (e.g., Baeyens et al., 1990; Wardle et al., 2007; Hütter et al., 2012), and relatedly, whether attention for the pairings or processing goals play a role (e.g., Dedonder et al., 2010; Prescott & Murphy, 2009). However, virtually all of these studies focused on the processes underlying EC effects. If the conditions for a process are not given, no EC effect will emerge; for example, if a memory account of EC is correct (Gast, 2018), there should be no EC effect *when* people do not remember the co-occurrence. Similarly, if a propositional account of EC is correct (Mitchell et al., 2009), then a ‘CS does not relate to US’ proposition would predict *when* EC effects should not occur (e.g., Unkelbach & Fiedler, 2016).

Thus, one can reframe the process question, the how question, as a *when* question: Manipulating the sufficient constituents for the *how* question (e.g., memory) answers *when* EC should occur and when not. In this sense, previous research provided answers to the *when* question, and we will address some of these answers in the general discussion. To fully understand the EC phenomenon, we nevertheless believe it is conceptually relevant to treat the *when* and the *how* questions independently.

4.1.1 The When Question

To delineate the *when* question more clearly, we need to explicate some assumptions underlying EC effects. According to the EC definition by De Houwer (2007), EC effects follow from the pairing of two stimuli, that is, the contiguity of CS and US. In the EC literature, however, the definition of contiguity is usually limited to “spatio-temporal pairing[s]” (De Houwer et al., 2020, p. 4). Thus, the construct of contiguity answers the *when* question from a behavioristic perspective, relying on the observable co-occurrence of CS and US in space and time (i.e., the environment) as the sufficient predeterminant of EC effects (e.g., Hütter et al., 2012, p. 599; Jones et al., 2010, p. 208). This behavioristic perspective affords a clear definition and high experimental control. Most process explanations (e.g., associative or propositional models) started from this behavioristic perspective (see Gast & De Houwer, 2012; De Houwer & Hughes, 2016; for exceptions) as the necessary condition.

To separate the *when* from the *how* on the cognitive side, one needs thus a cognitive construct that is influenced by spatio-temporal contiguity but does not directly implicate the underlying process (e.g., associations or propositions). We want to suggest the cognitive construct of the *mental episode* (Fiedler et al., 2005), which is functionally similar to an *event model* (Radvansky & Zacks, 2017) or *event file* (Hommel, 2004). In fact, Hommel and Stevenson (2021) recently tested predictions of such an event file account of attitudes. The basic idea is that memory organizes the stream of events around the organism into distinct episodes or events that guide future actions and help to develop expectations (Radvansky, 2012). For example, Fiedler et al. (2005) showed that, while keeping the space and time parameters constant, priming effects (e.g., detecting aggressive stimuli following a picture showing an aggressive act) vanished when prime and target were not encoded into the same mental episode. Similarly, Radvansky, Tamplin, and Krawietz (2010) showed that people were more likely to forget what they just put down on a table when they crossed a symbolic event boundary (i.e., leaving a room) compared to when they walked the same distance within the same room.

Many EC experiments, inadvertently or by design, employ principles that foster including CS and US in the same mental episodes. For example, one principle is Gestalt Psychology’s law of proximity: Two stimuli that appear close together are mentally grouped together (Wertheimer, 1938). Virtually all EC experiments utilize this law and present US and CS in close proximity. The spatial proximity is supported by the temporal configuration of trace conditioning trials in EC (i.e., 'CS – US – next trial'), which mirrors the auditory organization described by Wertheimer: "tap-tap, pause, tap-tap, pause" (p. 74[8]). Following such a rhythmic pattern is almost inevitable for most people (Neisser, 2014) and forces the CS and US into the same mental episode.

Similarly, procedural descriptions of EC experiments almost exclusively focus on the spatial and temporal parameters of the CS and US presentations (e.g., CS-US presentation or reinforcement schedules; e.g., Kattner et al., 2012; Kattner, 2014). To be sure, we do not claim that the organism was neglected in classic learning theory (e. g., Hull's construct of the organism's need as a general force or Seligman's idea of biological preparedness). Also, as already mentioned, in answering EC's *how* question, there are many examples of the organism's passive or active role in the learning process; for example, the role of memory or attention as predictors for EC effects (e.g., Corneille & Stahl, 2019; Stahl et al., 2009).

However, regarding the *when* question, the focus has been on the bottom-up, stimulus-driven input, while the top-down processes on the organism's side did not receive much consideration, as evinced by most EC experiments' procedural descriptions. To answer the *when* question, we assume that contiguity as the bottom-up, stimulus-driven component needs to be supplemented by a top-down, organismic-driven force. Together, these bottom-up and top-down forces determine the formation of mental episodes.

In the following, we introduce Fiedler and Bless's (Fiedler, 2001; Fiedler & Bless, 2000) dual-force model as a framework to predict whether two stimuli are encoded into the same mental episode and, thus, *when* a US influences a CS evaluation.

[8] Please note that in the original paper, we mistakenly indicated p. 303 as the source of the citation. However, the citation can be found on p. 74.

4.1.2 Determinants of Mental Episodes: A Dual-Force Perspective

Drawing from the work of Piaget (1954), the dual-force model (Fiedler, 2001; Fiedler & Bless, 2000) assumes two forces: An accommodative force that represents the external constraints of the environment and an assimilative force that imposes the organism's internal mental structures upon the external physical world. Both forces work simultaneously, yet their respective influence may vary. In most experimental paradigms, one may think of the accommodative force as the bottom-up, stimulus-driven influence and the assimilative force as the top-down, organism-driven influence. This distinction is readily applicable to EC research and the question of *when* people learn evaluations from co-occurrences.

As delineated above, the bottom-up influences are deeply embedded in EC research due to the origins of EC in the learning literature (Levey & Martin, 1975; Staats & Staats, 1958). This emphasis is still present today. For example, De Houwer et al. (2013) defined learning as "changes in the behavior of an organism that are the result of regularities in the environment of that organism" (p. 633). Notably, this definition's causal elements are the environment's regularities, while the organism is absent as a causal factor.

The top-down influences are the domain of social psychological theorizing, which stresses the person's active role, and whose intentional or unintentional construal of the social environment determines behaviors. For example, behavior may differ depending on what is active in people's memory (i.e., priming effects), depending on what people currently feel (i.e., mood effects), or depending on people's motivational states (i.e., self-regulation, goal setting effects). These processes may change and transform the input from the environment. In most scenarios, the two forces interact, but one may independently manipulate one or the other's strength.

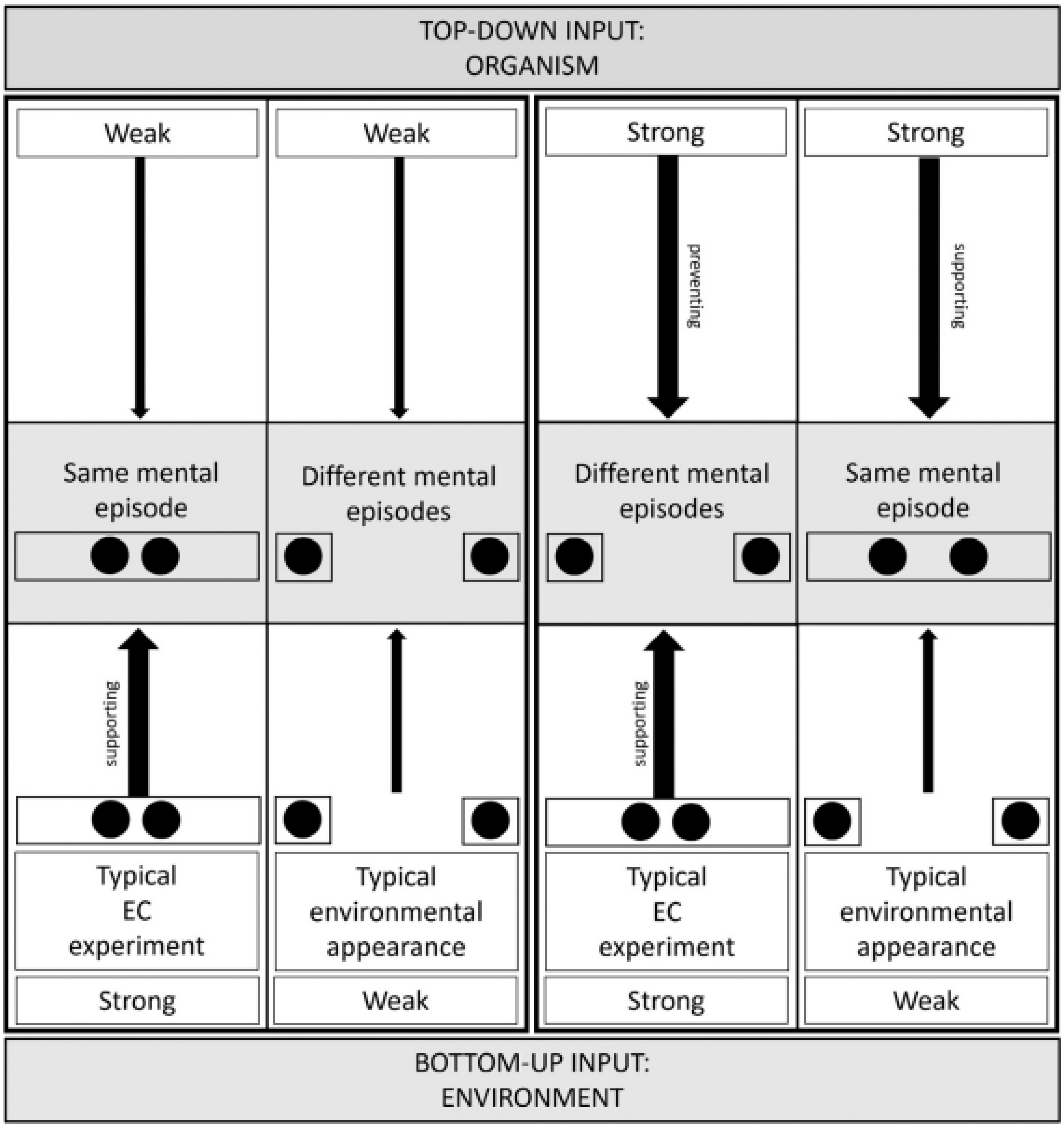
We treat regularities in the environment as bottom-up influences and the transformative influences upon this input within the organism as top-down influences. Thus, spatial-temporal contiguity is a bottom-up influence that forces CS and US into the same mental episode, while the law of proximity (Wertheimer, 1938) is a weak top-down influence; we will present examples of strong top-down influences below.

Figure 4.1 illustrates the two forces' interactions. The leftmost panel shows a strong bottom-up force (e.g., CS and US's repeated co-occurrence in time and space) with only weak input from the organism (e.g., Gestalt principles and perceptual organization), leading to the encoding of CS and US into the same mental episode. However, if the bottom-up influence is weak (e.g., single co-occurrences without spatial or temporal organization), illustrated by the larger spatial distance between dots (i.e., stimuli), CS and US will not be encoded into the same mental episode. This panel illustrates the case for many co-occurrences in people's daily lives that do not lead to EC effects without additional top-down input from the organism.

Figure 4.1's left half illustrates the typical behavioristic view in which the organism is a relatively passive recipient of environmental input. The 'weak' label on the organism's side highlights this aspect. This weak input could be Gestalt principles (see above) or a passive readiness for certain CS-US co-occurrences, akin to Seligman's (1970) preparedness idea (see below). Bruner's (1957) idea of perceptual readiness or existing mental structures such as schemas and stereotypes are further examples of weak top-down input. The right half represents the so far atypical cases in EC research. Even given a consistent and regular bottom-up influence, top-down influences might force CS and US into different mental episodes (e.g., Fiedler et al., 2005). Finally, even if the bottom-up influence is weak, strong top-down influences might force CS and US into the same mental episode (or event, or event file).

Figure 4.1

Illustration of Top-Down and Bottom-Up Influences on Mental Episodes.



Note. The size of the arrows represents the respective force's strength. The black dots represent CSs and USs. The first panel (from left to right) represents a standard EC-Paradigm; if spatio-temporal contiguity is high, CS and US are in the same mental episode. The second panel represents most stimuli in the world when both environmental as well as top-down influences are weak. The third panel shows the atypical case when environmental input is strong, but top-down processes force CS and US into separate mental episodes. The fourth panel illustrates the atypical situation when strong organismic input forces CS and US into the same mental episode, with no or weak environmental input.

Importantly, the representation of CS and US in the same mental episode constitutes only the necessary condition for EC effects to emerge. However, CS and US being in the same mental episode is not a sufficient condition for an EC effect. This point differentiates the *when* and the *how* question: we aim to answer *when* CS and US are represented in the same mental episode. If there is a transfer of valence and *how* this transfer operates is part of the *how* question. Thus, encoding two stimuli into the same mental episode still allows for the absence of EC effects and different processes to take place between the two stimuli.

Figure 4.1 should therefore be seen as independent of process assumptions about learning per se. One might argue that the top-down component immediately implies a propositional route to EC processes. This assumption would only be valid if any input from the organism's side is considered propositional. However, there are many examples from the literature that even in animals, organismic top-down components are highly relevant. Probably the most famous example of such organismic components is Garcia and Koelling's (1966) finding that rats learned to avoid saccharine taste after becoming ill (due to injections or radiation) or being shocked after drinking water laced with saccharine. However, they did not learn to avoid a bright light or clicking sound that was also presented when they touched the drinking spout of the water dispenser. Thus, from three spatio-temporal contiguities, rats learned an avoidance reaction only from one cue (see also Garcia, McGowan, Ervin, & Koelling, 1968). Within humans, a famous example is Seligman's (1970) preparedness hypothesis as an explanation of phobic reactions. Accordingly, species acquire a predisposition to learn some 'prepared' associations better than other 'non-prepared' or 'contraprepared' asssociations. Very few or a single learning trial may suffice for humans to acquire a fear reaction if the CS is represented as an organismically prepared stimulus. In addition, the spatial and temporal Gestalt principles by Wertheimer (1938) we discussed above are clearly on the organism's side but hardly propositional.

To illustrate the theoretical independence of the *when* from the *how* question, let us consider the two cases in Figure 4.1 that should lead to mental episodes. If CS and US frequently occur together with weak input from the organismic side (leftmost panel), the *when* question is answered affirmatively. However, the *how* answer could be associative (e.g., as in the surveillance paradigm; Olson & Fazio, 2001; but see Moran et al., 2021) and/or propositional (e.g., as with relational qualifiers; Hughes, Ye, Van Dessel, & De Houwer, 2019). This setup constitutes the most frequently discussed case in the literature. More interesting is the rightmost panel, which seems to presuppose a propositional process. However, if one considers organismic forces such as preparedness, schemas, or stereotypes, it is apparent that also the rightmost panel theoretically allows for associative and propositional answers. For example, in advertising, cruise ships (CS) may appear together with colorful sunsets (US), diapers (CS) may appear together with attractive parents (US), or expensive watches (CS) may appear together with celebrities (US). Therefore, the bottom-up force might meet a strong top-down force due to pre-existing concepts and schemas. While the observer may even actively try to avoid the influence of the co-occurrence (i.e., it is blatant advertisement), and it may be clear that the co-occurrence is propositionally invalid (i.e., parents changing diapers do not look so happy), EC effects may nevertheless occur due to other processes. Practically, experimental manipulations might imply both the *when* and the *how* questions, but theoretically, they should be considered independent.

In the following, we establish the independence with experimental configurations that should lead to EC effects from both process explanations of EC but not from the present dual-force perspective. We will focus on propositional and associative explanations; however, our arguments also apply to other process explanations (memory-based, Gast, 2018; misattribution, Jones et al., 2009).

4.1.3 The Present Research

We will present a test of the dual-force model applied to EC. The most interesting test based on Figure 4.1 is a situation in which the spatio-temporal proximity of CS and US should lead to an EC effect, both from an associative and propositional perspective, but top-down influences force CS and

US into separate mental episodes. Such a result would be difficult to explain without considering the *when* question.

To create the experimental situation, we borrowed the setup of Fiedler et al. (2005, Experiment 1). They showed that one might facilitate word recognition by presenting associated pictures; for example, showing a man stabbing a woman facilitates recognizing the word ‘aggressive’. If participants provided a judgment about the picture, for example, ‘Is this behavior aggressive?’, the priming advantage was significantly reduced. The authors argued that this follows because judging the behavior closes the mental episode, and priming effects necessitate an open mental episode in which the prime (i.e., the picture) and the target (i.e., the word) fall into the same episode. In other words, the top-down judgment disrupted the bottom-up influence and the typical judgment facilitation for the word stimuli (i.e., priming effect).

As alluded to above, Radvansky (2012) made a similar observation. According to his *Event Horizon Model*, “[…] crossing an event boundary, such as from one location to another, can disrupt memory, even memory for information that continues to be relevant” (p. 269). As event boundaries do not just refer to locations but also to breaks in causal chains or the inception of a new activity, a top-down judgment process should lead to the segmentation of the stream of consciousness and interfere with the bottom-up force. Similarly, Shin and DuBrow (2021) noted that people segment continuous information into distinct events with a defined beginning and end. When an event ends, either due to bottom-up (e.g., perceptual) or top-down (e.g., inference) processes, the information within this segment fades as it is no longer relevant.

In the following, we apply this logic to EC with bottom-up and top-down configural constituents that establish or disrupt mental episodes. Concretely, we will use a sequential backward conditioning paradigm, in which the temporal configuration of (‘US – CS – pause’) strongly resembles the *Gute Gestalt* of tapping described by Wertheimer (‘tap – tap – pause’; 1938). This configuration should force US and CS into the same mental episode as long as the top-down input is weak. The setup is also akin to the memory model by Radvansky (2012), according to which the organism segments the stream of events in the environment into mental episodes.

The inclusion of a top-down judgment provides another force. We will include a judgment of the US before the CS, and this top-down influence should interrupt the Gestalt of the trial (i.e., 'US – judgment – CS – pause') and close the mental episode after the US (Fiedler et al., 2005). The judgment in-between US and CS should force US and CS into separate mental episodes despite their spatio-temporal proximity. Accordingly, we predict a reduction of the EC effect under these conditions (Experiments 1-3). Additionally, we will show that another top-down process may force US and CS into the same mental episode (Experiment 4) and that the same manipulation works for forward conditioning (i.e., 'CS – US – pause'; Experiment 5). Finally, we will show that the manipulation works when the bottom-up input is identical between conditions (Experiment 6).

In summary, we will hold spatio-temporal proximity constant but manipulate the top-down influence force (Fiedler et al., 2005; Radvansky, 2012) to encode CS and US in the same or different mental episodes leading to the presence or absence/reduction of EC effects, respectively. The absence of EC effects follows if the separation into different mental episodes is successful on every learning trial; reduction follows if separation is successful on some but not all trials compared to the no judgment conditions (cf. Figure 4.2).

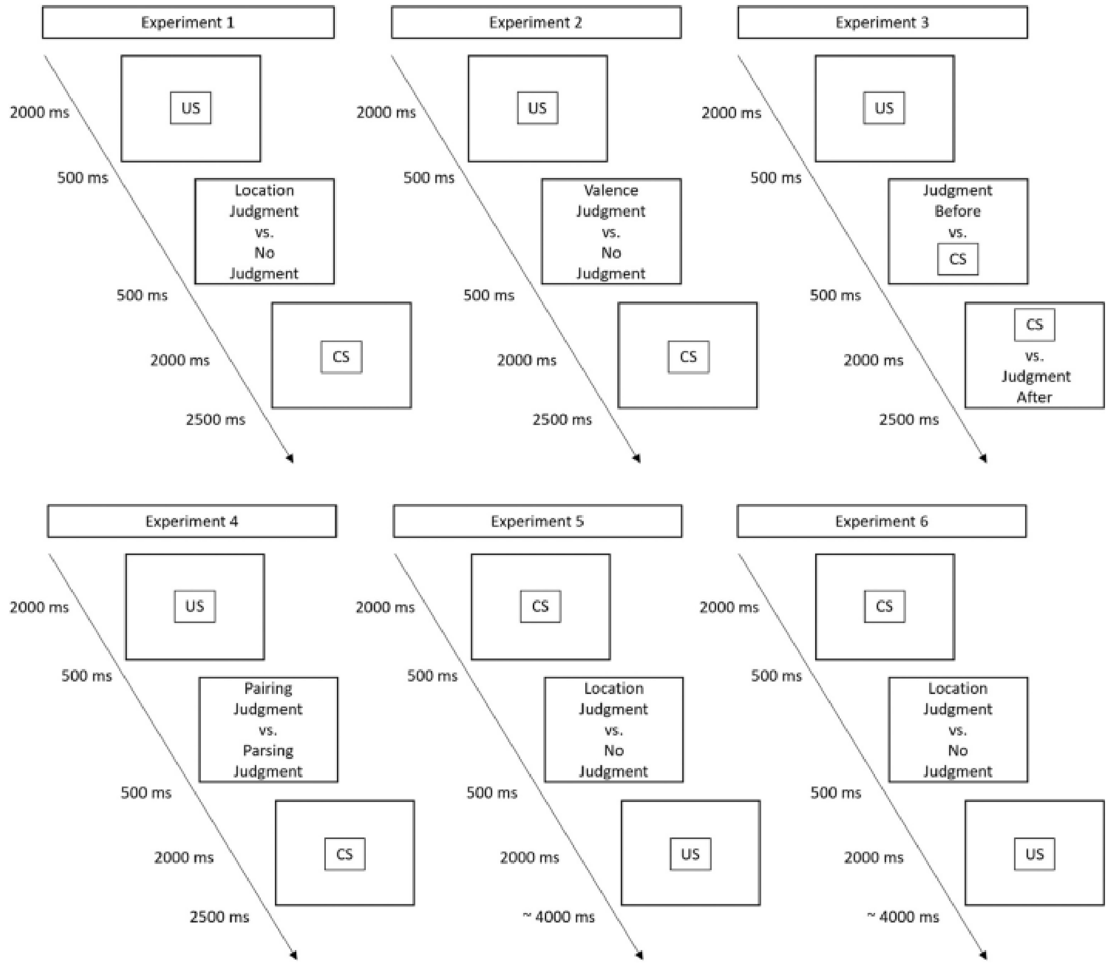
4.2 Preview of Experiments

We conducted nine experiments in this research line, of which we will report six. Below we provide a summary of the three remaining experiments, which are also described in the supplements.[9] Figure 4.2 illustrates the design of these six experiments. The first four experiments used a backward conditioning procedure (cf. Kim et al., 2016; Martin & Levey, 1978; Stuart et al., 1987), showing the US first and the CS shortly afterward.

[9] The present Experiment 1 was the third in the series and Experiment 2 the second. The first, fifth and sixth experiment in the series are presented in the supplement.

Figure 4.2

Overview of the Conditioning Trials in Experiments 1-6.



Note. The CS appears after the US (backward conditioning procedure) in Experiments 1-4. The location/valence judgment had participants answer a question about the previously presented US's location/valence. In Experiments 5-6, the CS appears before the US (forward conditioning procedure). The location judgment had participants answer a question about the previously presented CS's location. The 'no judgment' condition showed a blank screen in-between stimulus presentations (Experiments 1, 2, 4, & 5). In Experiment 3, we manipulated the timing of the judgment (before vs. after CS presentation). Thus, participants in the 'judgment before' condition had to make a valence judgment about the US before the CS appeared; participants in the 'judgment after' condition had to answer the valence question after US and CS presentations. In Experiment 6, the blank screen in the 'no judgment' condition contained grey bars to keep the bottom-up input between conditions comparable.

We used this backward conditioning procedure to allow participants to make a judgment about the US, thereby ensuring the processing of the US. The stimulus presentation order should not influence the EC effect (Gast, Langer, & Sengewald, 2016). The last two experiments used a forward conditioning procedure, showing the CS first and the US shortly afterwards. We always presented CS and US in close temporal (i.e., 2000 ms lag in-between presentations) and spatial proximity (i.e., at the same screen location; except for Experiment 1, when the CS appeared in the middle of the screen, and the US appeared either above or below the later CS location). Thus, we aimed to keep the bottom-up force constant across experiments.

Experiment 1 presented an initial test of the dual-force model. We asked participants to judge the location of the US before the CS was presented. Following Fiedler et al.'s (2005) procedures, the judgment should segment US and CS into different mental episodes, thereby preventing EC effects.

In Experiment 2, participants judged the US's valence instead of the US location. From a propositional perspective, one should expect that the re-affirmation of stimulus valence should strengthen EC effects. From a dual-force perspective, though, we predicted that the judgment still forces US and CS into different mental episodes, again preventing EC effects.

Experiment 3 then highlights the critical role of the mental episode construct. It should not be the judgment per se during learning, but the judgment in-between the US and CS that forces US and CS into different mental episodes. Thus, we predicted that the EC effect vanishes when participants make a judgment in-between the US-CS presentation but not when participants make the same judgment after the US-CS presentation. Of relevance, the latter condition's mental load is higher than the former (i.e., participants must keep US valence in working memory).

Experiment 4 then shows that EC effects re-appear if one incorporates another top-down process, highlighting that US and CS belong into the same mental episode. Experiment 5 replicates Experiment 1 using a forward conditioning procedure. Experiment 6 addresses a final concern. The presentation of the judgment prompt could be seen as a bottom-up rather than a top-down influence. Experiment 6 provides a comparable bottom-up input in both conditions, varying only the topdown component (i.e., the judgment).

As stated, we do not report three additional experiments conducted in this line of research. The first of these three experiments compared valence and location judgments. Both judgment types successfully eliminated the EC effects; yet, they are fully redundant with Experiments 1 and 2 but lacked proper control conditions. The second experiment also used an anagram reasoning task in-between US and CS presentation to establish relations between CS and US; however, this subtle manipulation failed to re-establish the EC effect. The third experiment simultaneously implemented four changes: (a) we used forward instead of backward conditioning, (b) we asked participants to make a judgment after the CS-US presentation, (c) we showed scrambled text in-between CS-US presentation in the 'judgment after' condition, and (d) we conducted the experiment online instead of the laboratory-based. Even in the control condition, the experiment did not show significant EC effects. However, it was not transparent which of these changes led to the insignificant findings in this experiment. We thus conducted two additional experiments, implementing only one design change at a time (Experiments 5 and 6) and successfully replicated our findings. We discuss potential explanations in the supplements, where we report these three experiments.

4.3 Transparency and Open Science

We pre-registered the hypotheses for Experiments 1 to 4 based on a priori theorizing. We pre-registered Experiments 5 and 6 to address reviewer concerns. We based the sample sizes for the first four experiments on the average EC effect size reported in the meta-analysis by Hofmann et al. (2010), assuming a standard EC effect (Simonsohn, 2014). Due to the COVID-19 pandemic, we conducted the two new experiments online, using new US material (for details, see Experiments 5 and 6). As we did not know how these changes would influence the effect size, we used sequential analysis (Lakens, 2014), assuming that we may already detect an existing effect collecting data from 150 participants but keeping open the possibility to increase the sample size to 300 if necessary with adjusted *p*-values. We report all data exclusions, which were all pre-registered, all manipulations, and all measures. The ‘Open Practices’ section below provides the respective links. We conducted this research according to the American Psychological Association’s (2017) Ethical Principles in the Conduct of Research with Human Participants.

Please note that we pre-registered a nested contrast analysis and predicted a significant EC effect vs. the absence of an EC effect in the respective conditions. Instead of an interaction, this analysis has more power for the predicted attenuation pattern compared to the ANOVA interaction (cf. Schad et al., 2020). As most readers are more familiar with the ANOVA results, we will report the ANOVA outcomes. We report the nested contrast analysis in the supplementary materials. Except for Experiment 6, both analytic strategies lead to the same inferential conclusions. The pre-registered secondary regression analysis (Experiment 1-4) supports our findings as well and is also available from the supplements. Additionally, we will report results from a Bayesian paired samples t-test for conditions for which we predicted and found a reduced EC effect. For our predictions, the interesting test is whether there is more evidence for H0 (i.e., there is no difference between CS ratings) than for H1 (i.e., there is a difference between CS ratings). Thus, we will report the inverse Bayes Factor (i.e., BF01 or 1/ BF10). Our interpretation of the evidence for H0 follows the criteria set by Lee and Wagenmakers (2013): anecdotal (i.e., not enough evidence to support or reject H0) = 1-3, moderate = 3-10, strong = 10-30, very strong = 30-100, extreme >100.

We want to highlight that the Bayesian analysis does rely on standard priors as we had no a priori effect size estimations. The pre-registered comparison of the EC effect with a control condition in the form of a nested contrast or interaction as described above provides a stronger test, as it relies on the presence of an effect, not its absence.

4.4 Experiment 1: Forcing CS and US into Separate Mental Episodes

Experiment 1 investigated whether we can force US and CS into separate mental episodes. To do so, we asked participants to make judgments in-between stimulus presentations while holding the spatio-temporal proximity constant. This method is a direct adaption of the prior work of Fiedler et al. (2005) for semantic priming applied to EC. An a priori power analysis using the G*Power software

(Faul et al., 2007) showed that 32 participants should suffice to detect a common EC effect (d = 0.52 with a correlation of r = 0.50 between DVs; Hofmann et al., 2010), with α = 0.05 and 1-β = 0.80. Because we aimed to moderate this effect, we more than quadrupled this number (see Simonsohn, 2014) and pre-registered a sample size of 150 participants. Additionally, we conducted sensitivity power analyses for all experiments, again using the G*Power software (Faul et al., 2007), which we report in the method section of every experiment.

4.4.1 Method

Participants. As pre-registered, we collected data from 150 participants on a University campus who participated for payment of 2 Euro (about $2.30). Also, as pre-registered, we excluded eleven participants because they reported low concentration and one further participant due to reporting speaking or understanding Japanese or Chinese, as we used Kanjis as CSs. Thus, we retained 138 participants in the sample (108 women, 30 men; M_{age} = 22.90, SD_{age} = 4.13). A mixed ANOVA with 69 participants per condition (N = 138) would be sensitive to effects of ηp 2 = 0.06 with 80% power (α = 0.05, r = 0.50). This means the experiment would not be able to detect effects smaller than ηp^2 = 0.06 reliably.[10]

Materials. As USs, we used 60 pictures from Förderer and Unkelbach (2012) showing various animals. Participants rated these, and we selected the idiosyncratically best- and least-liked stimuli as USs. As CSs, we used eight Kanjis (i.e., Japanese characters). We programmed the experiment with OpenSesame (Mathôt et al., 2012).

Design. The computer program assigned participants alternatingly to either the ‘location judgment’ condition or the ‘no judgment’ condition (see Figure 4.2). We varied US valence within participants, resulting in a 2 (‘no judgment’ vs. ‘location judgment’; between) x 2 (US valence: ‘positive’ vs. ‘negative’, within) mixed design.

[10] The sensitivity power analysis outputs f as effect size. Because we used ηp^2 (i.e., partial eta squared) as effect size in our main analysis (i.e., mixed ANOVA), we converted the f value calculated by G*Power (as recommended Cohen, 1988) into ηp^2 using the transformation option in G*Power’s a priori power analysis type.

Procedure. Upon arrival, experimenters accompanied participants to individual PC workstations and started the OpenSesame program. After providing informed consent, participants pre-rated how much they liked 60 pictures on a 5-point-scale ranging from − 2 to +2 with three labels (‘not at all’, ‘neutral’, ‘very much’) to select the USs and calibrate participants to the rating scale (Unkelbach et al., 2012). The supplements present the specific instructions.

The program selected the four most-liked and the four least-liked pictures as USs. Then the conditioning phase started. The program randomly assigned the eight CSs to the eight USs, and each CS was assigned to only one US (i.e., ‘one-to-one’ assignment; see Stahl & Unkelbach, 2009). A US appeared eight times with its assigned CS in both conditions, resulting in 64 trials. The program randomized the trial order.

In the ‘location judgment’ condition, instructions informed participants that they would need to judge the pictures’ location using two keys, specifically, whether a picture was in the upper or lower half of the screen. A US appeared in the lower or upper half of the screen for 2000 ms in a given trial. The program counterbalanced location across trials and USs. When the US disappeared, a black screen followed for 500 ms, and then a screen with the location question appeared. Participants used the ‘A’ (upper half) and ‘L’ (lower half) keys to judge the location. After the keypress, the CS appeared with a delay of 500 ms in the screen center and stayed onscreen for 2000 ms. Then the CS disappeared, and after 2500 ms of a black screen, the next trial started.

In the ‘no judgment’ condition, instructions informed participants that they would need to observe the presented pictures attentively. A given trial was similar to the judgment condition; however, after the US disappeared, participants only saw a black screen for 1900 ms (i.e., 500 ms plus 900 ms plus 500 ms; cf. Figure 4.2), and then the CS appeared. We selected the time based on our experience with how much time people need to press a key to judge stimulus locations.

After the 64 conditioning trials, the CS rating phase started. The program instructed participants that their task would be to rate their liking of Chinese characters. For a given rating, each

CS was presented once in the screen's center with a 5-point-scale beneath it, ranging from − 2 to +2 with three labels ('not at all', 'neutral', 'very much'). The program randomized the rating order for each participant. Upon completing the rating phase, participants provided demographic information, and experimenters debriefed and paid them.

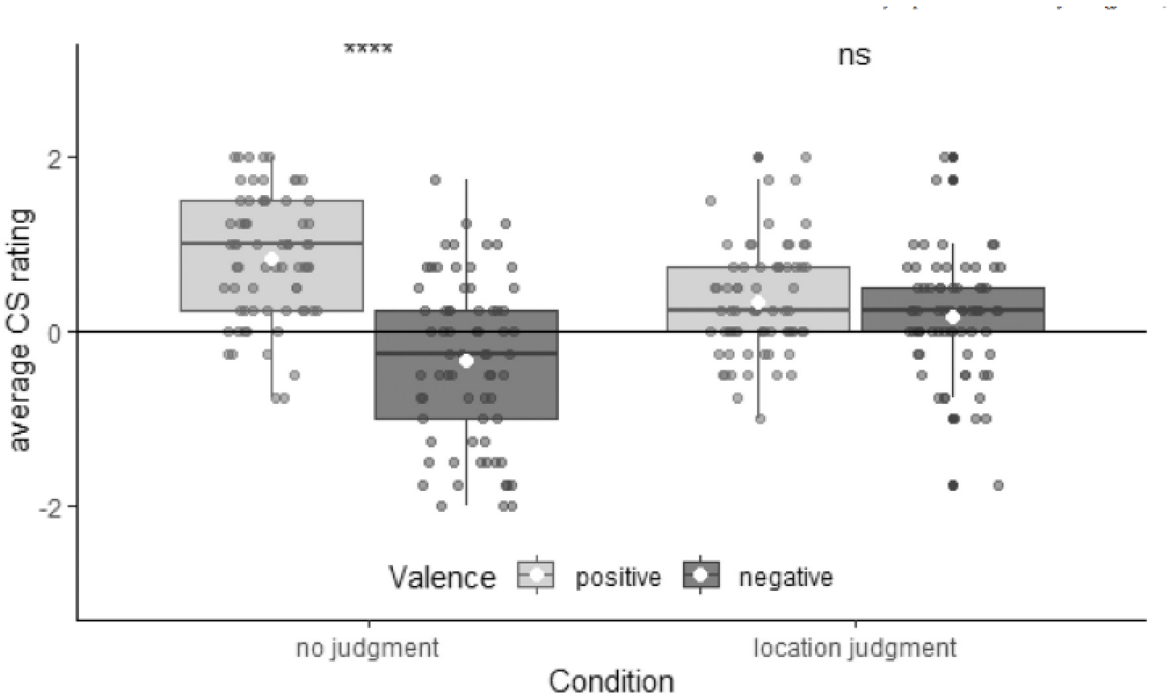
4.4.2 Results

Location Judgments. In the 'location judgment' condition, participants, on average, pressed a key within 843 ms (*SD* = 289), which was not significantly different from the 900 ms delay we implemented in the 'no judgment' condition, $t(68) = -1.62$, $p = .110$, $d = 0.19$; thus, the timing conditions were roughly comparable between conditions. Participants classified 98% of USs correctly; that is, participants misclassified on average 1 of 64 USs.

CS Ratings. Figure 4.4 presents participants' CS ratings averaged within US valence in the 'no judgment' and the 'location judgment' conditions. We analyzed these data with a condition (no judgment vs. location judgment) by US valence (positive vs. negative) mixed ANOVA with repeated measures on the latter factor. This ANOVA showed a valence main effect, $F(1, 136) = 44.58$, $p < .001$, $\eta p^2 = 0.25$, replicating a typical EC effect. On average, participants rated CSs presented with positive USs as more likeable ($M = 0.59$, $SD = 0.71$) than CSs presented with negative USs ($M = -0.07$, $SD = 0.84$). Importantly, as Figure 4.4 suggests, this main effect was moderated by the judgment conditions, $F(1, 136) = 25.12$, $p < .001$, $\eta p^2 = 0.16$. In the 'no judgment' condition, participants rated CSs presented with positive USs significantly more positive ($M = 0.85$, $SD = 0.71$) than CSs presented with negative USs ($M = -0.32$, $SD = 0.95$), $t(68) = 6.66$, $M_{diff} = 1.17$, 95%CI [0.82, 1.52], $p < .001$, $d = 0.81$. As the interaction indicates, this difference in ratings was significantly smaller in the 'location judgment' condition, where the difference was no longer significant: Participants rated CSs presented with positive US only slightly more positive ($M = 0.34$, $SD = 0.62$) than CSs presented with negative USs ($M = 0.17$, $SD = 0.64$), $t(68) = 1.73$, $M_{diff} = 0.17$, 95%CI [− 0.03; 0.36], $p = .087$, $d = 0.21$. Finally, the ANOVA main effect for condition was not significant, $F(1, 136) < 0.01$, $p = .943$, $\eta p^2 < 0.01$.

Figure 4.4

Combined Box and Scatter plots of Participants' CS Ratings in Experiment 1 as a Function of US Valence (Within Participants) and Experimental Condition ('No Judgment' vs. 'Location Judgment'; Between Participants), Illustrating the Central Interaction.



Note. The simple effects for valence show that the difference between CS ratings in the 'no judgment' condition is highly significant but not significant in the 'location judgment' condition. Higher values indicate more positive evaluations. The box plots constitute the first quartile, third quartile, and median. The circle within a box plot indicates the mean. Each dot represents an evaluation. Observations outside the whiskers are outliers with >150% interquartile range distance from the next quartile.

4.4.3 Bayesian Analysis

We predicted a reduced EC effect in the 'location judgment' condition, which, as expected, yielded a non-significant result. To test whether this finding provides evidence for a null effect, we additionally computed a Bayes factor analysis with Jeffrey's Amazing Statistics Program (JASP). As we are interested in the evidence for H0 (compared to H1), we will report the inverse Bayes Factor. Using a JZS prior of 0.707[11], the analysis showed that BF01 = 1.83, which can be interpreted as anecdotal evidence (Lee & Wagenmakers, 2013) for H0 (i.e., there is no difference between measures). Thus, our data is almost twice as likely under the hypothesis that there is no effect as under the hypothesis that there is an effect (cf. Rouder et al., 2009).

4.4.4 Discussion

Experiment 1 showed a typical EC effect; however, the location judgment in-between USs and CSs significantly reduced the EC effect to a non-significant difference. Thus, while the spatio-temporal proximity between CSs and USs was identical, the additional judgment all but eliminated the EC effect. We predicted this moderation based on the assumption that judging the US forces US and CS into separate mental episodes, the same way that judgments between primes and targets prevent priming effects (Fiedler et al., 2005).

Thus, Experiment 1 provides first evidence for a dual-force perspective on EC and a potential answer to the *when* question in EC. However, before we interpret these findings in more depth, we must consider alternative explanations; that is, factors that also varied between conditions beyond the judgment tasks.

First, the present setup may lead to confusion due to the backward conditioning procedure; participants could have understood the task as a forward conditioning task. However, the analyses of the CSs as a function of the subsequent USs in this and the following experiments did not show standard EC effects. Forward conditioning as an alternative explanation can thus be eliminated. The analyses, R codes, and result reports are available on our OSF page (https://osf.io/57ux2/?view_only=b5f058da184a4d24b2a951f44ef30ab9).

Second, we controlled the timing between US presentation and CS onset in the backward conditioning procedure in the 'no judgment' condition, while participants controlled it in the 'location

[11] This Jeffery-Zellner-Siow prior (r = 0.707) is the current default of the BayesFactor package (Morey et al., 2015) in R and in JASP (cf. Dienes, 2019).

judgment' condition. However, the average difference between these conditions was not significant, and in fact, the 'location judgment' condition had an overall shorter delay between US presentation and CS onset.

Third, and most critically, participants may not have encoded US valence during US presentation. They might have solved the judgment task without encoding US valence, and despite the attention-grabbing power of evaluative information (e.g., Pratto & John, 1991; see Unkelbach et al., 2020; for a review), participants might not have paid attention to the pictures at all, or they might have focused on different attributes of these pictures (Förderer & Unkelbach, 2015). If participants did not attend to US valence, a reduction of the EC follows as well (see Gast & Rothermund, 2011a). Experiment 2 directly addresses this possibility.

4.5 Experiment 2: Focusing on Valence

Experiment 2 investigated whether a lacking valence focus might be responsible for the lack of an EC effect in the 'location judgment' condition. To address this possibility, we changed the judgment task from a location judgment to a valence judgment; that is, in Experiment 2, participants evaluated US valence in between the US and CS presentations. In addition, for most models that answer the how question of EC, explicitly affirming US valence should strengthen the EC effect; however, we argue that the judgment process forces US and CS into separate mental episodes, preventing EC effects.

4.5.1 Method

Participants. We pre-registered to collect data from 150 participants but collected data from 151 participants, who participated for 2 Euro (about $2.30). Before analyses, we excluded one participant who did not finish the experiment. As pre-registered, we excluded eighteen additional participants who stated low concentration. Also, as pre-registered, we excluded six further participants who indicated to speak or understand Chinese or Japanese. Finally, we excluded two participants because they were younger than 18 years and could not provide informed consent for participation and open access to their data. We therefore retained 124 participants for analysis (66 women, 56 men, 2 other; M_{age} = 22.80, SD_{age} = 5.89). A mixed ANOVA with 61 participants in the ‘no judgment’ condition and 63 participants in the ‘valence judgment’ condition (N = 124) would be sensitive to effects of $\eta p^2 = 0.06$ with 80% power ($\alpha = 0.05$, $r = 0.50$). This means the experiment would not be able to reliably detect effects smaller than $\eta p^2 = 0.06$.

Design, materials, and procedure. Experiment 2 was highly similar to Experiment 1 (cf. Figure 4.2), except that the judgment condition asked participants to indicate whether the presented US picture in a given trial was positive or negative valence via two response keys. The program presented the US in the middle of the screen. Everything else was identical to Experiment 1.

4.5.2 Results

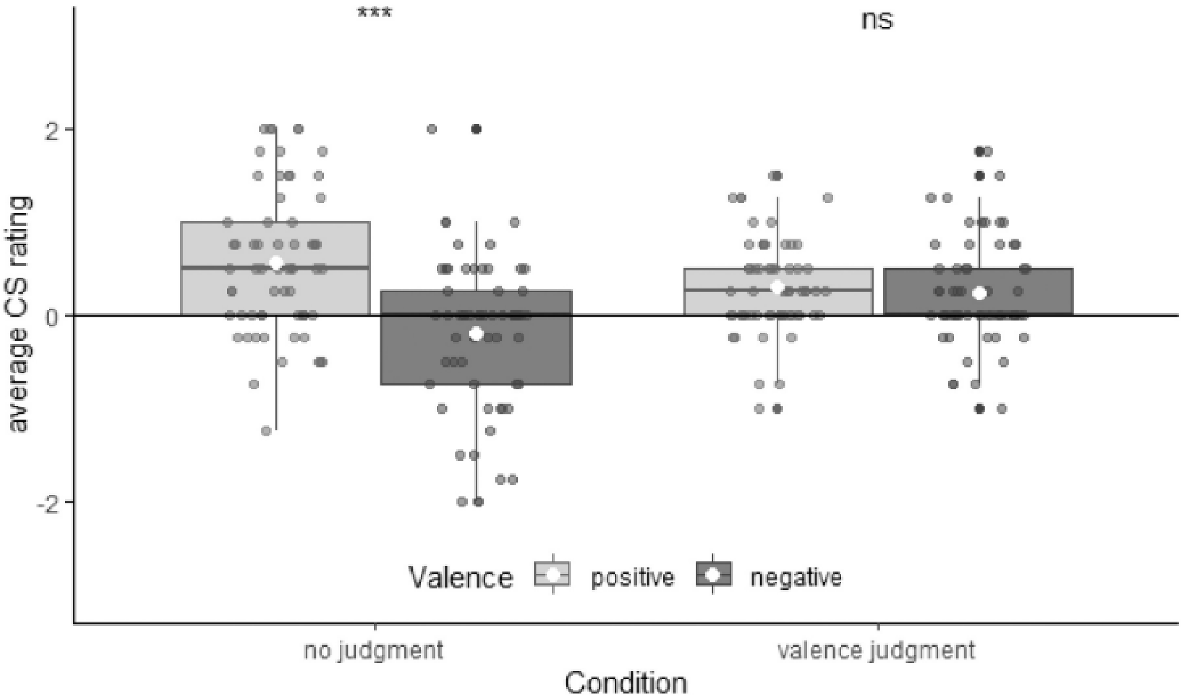
Valence Judgments. In the ‘valence judgment’ condition, participants, on average, pressed a key within 965 ms (SD = 444), which was not significantly different from the 900 ms delay we implemented in the ‘no judgment’ condition, $t(60) = 1.14$, $p = .259$, $d = 0.15$; thus, timing conditions were again roughly comparable between conditions. Participants classified 90.50% of the USs correctly; that is, participants misclassified on average 6 of 64 presented USs.

CS Ratings. Figure 4.5 presents participants’ CS ratings averaged within US valence in the ‘no judgment’ and the ‘valence judgment’ conditions.

We analyzed these data the same way as in Experiment 1. The ANOVA again showed a valence main effect, $F(1,122) = 17.17$, $p < .001$, $\eta p^2 = 0.12$, replicating a typical EC effect. On average, participants rated CSs presented with positive USs as more likeable ($M = 0.43$, $SD = 0.67$) than CSs presented with negative USs ($M = 0.01$, $SD = 0.73$). As Figure 4.5 suggests, this main effect was moderated by the judgment condition, $F(1, 122) = 12.16$, $p < .001$, $\eta p^2 = 0.09$. In the ‘no judgment’ condition, participants rated CSs presented with positive USs more positive ($M = 0.56$, $SD = 0.78$) than CSs presented with negative USs ($M = -0.20$, $SD = 0.81$), $t(62) = 4.28$, $M_{diff} = 0.76$, 95%CI [0.41, 1.12], $p < .001$, $d = 0.54$.

Figure 4.5

Combined Box and Scatter plots of Participants' CS Ratings in Experiment 2 as a Function of US Valence (Within Participants) and Experimental Condition ('No Judgment' vs. 'Valence Judgment'; Between Participants), Illustrating the Central Interaction.



Note. The simple effects for valence show that the difference between CS ratings in the 'no judgment' condition is highly significant but not significant in the 'valence judgment' condition. Higher values indicate more positive evaluations. The box plots constitute the first quartile, third quartile, and median. The circle within a box plot indicates the mean. Each dot represents an evaluation. Observations outside the whiskers are outliers with >150% interquartile range distance from the next quartile.

As the interaction indicates, this difference in ratings was again significantly smaller in the 'valence judgment' condition, where the difference was no longer significant: Participants rated CSs presented with positive USs only slightly more positive (M = 0.29, SD = 0.50) than CSs presented with negative USs (M = 0.23, SD = 0.57), $t(60) = 0.76$, $M_{diff} = 0.07$, 95%CI [− 0.11; 0.24], $p = .448$, $d = 0.10$. Finally, the ANOVA main effect for condition was not significant, $F(1, 122) = 1.32$, $p = .253$, $\eta p^2 = 0.01$.

4.5.3 Bayesian Analysis

In Experiment 2, the EC effect for the 'valence judgment' condition again was non-significant. Thus, we computed a Bayes factor analysis to provide further evidence for a null effect. Using the same program and procedure as in Experiment 1, the inverse Bayes factor analysis yielded BF01 = 5.40, which can be interpreted as moderate evidence (Lee & Wagenmakers, 2013).

Thus, the observation of our data is five times more likely under the assumption that there is no effect than under the assumption that there is an effect (cf. Rouder et al., 2009).

4.5.4 Discussion

Experiment 2 replicated Experiment 1's results. Despite a judgment task that necessitates the encoding and processing of US valence, the judgment in-between US and CS presentation significantly reduced the EC effect to non-significance. Experiment 2 thereby takes care of a central concern in Experiment 1, namely that participants solved the judgment task without processing US valence.

Participants' average response latency in the judgment task was again comparable to the no judgment condition's fixed interval. Unlike Experiment 1, participants took slightly longer than 900 ms to make the valence judgments. Together with Experiment 1, this pattern makes the differential timing of CS-US presentation an unlikely explanation for the reduced EC effect. In addition, given participants' classification rates, the absence of EC effects is unlikely due to misperceptions of US valence.

We want to emphasize that the valence judgment eliminated the EC effect, despite participants confirming that a 'good' (or 'bad') stimulus preceded the CS. Given that we find standard EC effects without the judgment, the results of Experiments 1 and 2 do not follow from EC accounts that answer the how question without substantial additional assumptions. One may argue that a misattribution account would predict this pattern; however, the misattribution explanation relies on simultaneous CS-US presentation (e.g., Hütter & Sweldens, 2013; Sweldens et al., 2010) and would not predict the EC effects in the present ‘no judgment’ conditions. Importantly, from a propositional perspective, affirming the US valence should foster propositions about valence and the CS.

However, if one considers the typical EC paradigm with the organismic influence in mind, for example, as an instance of a Gestalt experiment that implements a tap-tap-pause rhythm and the rule of what goes together belongs together (Wertheimer, 1938), then the judgment not only interrupts this pattern but also forces US and CS into different episodes and thereby prevents, in our parlance, the necessary condition for EC effects.

Still, one might argue that EC is a fragile phenomenon (i.e., an “ephemeral” effect; Jones et al., 2010, p. 209), and any additional task during learning might reduce the effect. This assumption would also answer the *when* question: EC may mainly happen in controlled laboratory settings that create exceptional circumstances as described in the literature. While possible, this argument would rob the EC phenomenon of its applied relevance. Given the replicability of the EC effect across labs and procedures, we deem this unlikely. However, less drastically, we might have violated another pre-requisite for EC effects, particularly in a sequential paradigm. If one assumes that EC effects depend on memory for the joint CS-US presentations (e.g., Gast, 2018; Gast et al., 2012; Stahl et al., 2009), then any task that reduces attention during the learning phase and subsequently memory should reduce EC effects.

One example for such a reduction in EC effects due to reduced attention is provided by Field and Moore (2005). They showed in two experiments that attention during learning is a relevant factor to establish EC effects, independent of memory for CS-US presentations. Pleyers et al., (2009) also manipulated attention during learning and found that reduced attention leads to reduced memory and reduced EC effects. Halbeisen and Walther (2015) also claimed that a secondary task interferes with EC, particularly when the secondary task shares the same stimuli. Similarly, Mierop et al. (2017) employed a secondary task during learning and found that the secondary task reduced the EC effect. Thus, secondary tasks during learning indeed diminish EC effects by influencing attention, memory, or both.

The present judgment task cannot be compared in terms of mental load to the demanding secondary tasks used by Halbeisen and Walther (2015) or Mierop et al. (2017).

Nevertheless, a secondary task's potentially disruptive influence must be addressed for our claim that the present moderation follows from forcing US and CS into different mental episodes. Thus, Experiment 3 manipulated the critical aspect of timing the judgment.

4.6 Experiment 3: The Role of Timing

The judgment tasks in Experiments 1 and 2 may present a 'mental load' during learning and thereby reduce attention and memory, leading to reduced EC effects. Experiment 3 addressed this possibility by changing the 'no judgment' condition into a condition with factually more load than Experiments 1 and 2's 'judgment' condition; however, without forcing US and CS into separate mental episodes. We achieved this by moving the US valence judgment from before the CS presentation (forcing a new mental episode) to after the CS presentation (leaving the US-CS episode intact). Thus, participants judged US valence during the learning phase in both conditions. The judgment after the CS presentation should incur a higher mental load. Participants need to keep the correct response active in working memory during CS presentations. We nevertheless expect an EC effect in this condition, as both US and CS should be encoded into the same mental episode.

In contrast, participants who judge US valence before the CS presentation can deliver their response with minimal delay; however, the judgment should force separate mental episodes. We thus expected to replicate the absence of an EC effect when participants judge US valence in-between US and CS but expected a standard EC effect when participants judge US valence in-between US-CS trials (cf. Figure 4.2).

4.6.1 Method

Participants. As pre-registered, we collected data from 150 participants on a University campus who participated for payment of 2 Euro (about $2.30). Also, as pre-registered, we excluded eighteen participants because they reported low concentration. One participant did not finish the experiment. Third, as pre-registered, we excluded six further participants who indicated speaking or understanding Chinese or Japanese. Finally, we excluded three participants because they were younger than 18 years and could not provide informed consent for participation and open-data access. We thus retained 122 participants for analysis (76 women, 44 men, 1 other, 1 prefer not to say; M_{age} = 24.40, SD_{age} = 8.20). A mixed ANOVA with 60 participants in the 'judgment after' condition and 62 participants in the 'judgment before' condition (N = 122) would be sensitive to effects of $\eta p^2 = 0.06$ with 80% power ($\alpha = 0.05$, $r = 0.50$). This means the experiment would not be able to reliably detect effects smaller than $\eta p^2 = 0.06$.

Design, materials, and procedure. Experiment 3 was highly similar to Experiment 2 (cf. Figure 4.2). The sole difference was that we changed the 'no judgment' condition to a 'judgment after' condition. Thus, the design was a 2 (valence judgment: before vs. after CS presentation; between) x 2 (US valence: positive vs. negative; within) mixed design. Participants in the 'judgment after' condition provided the same US valence judgment as participants in the 'judgment before' condition, but only after the CS disappeared (see Figure 4.2). To ensure that participants judged the US and not the CS in the 'judgment after' condition, we specifically asked for the valence of the animal picture. Everything else was identical to Experiments 1 and 2.

4.6.2 Results

Valence Judgments. In the 'judgment before' condition, participants on average pressed a key within 1010 ms (SD = 379 ms), which was not significantly different from the 'judgment after' condition (1101 ms, SD = 464 ms), $t(113.84) = -1.17$, $p = .246$, $d = 0.21$.[12] More importantly, timing between US and CS presentation did not differ between conditions. In the 'judgment before' condition, participants pressed the key within 1010 ms (as mentioned before). In the 'judgment after' condition, a blank was shown once for 1000 ms after the US presentation. There was no difference between timing, $t(61) = 0.22$, $p = .825$, $d = 0.03$. In addition, participants classified 91% of the US

[12] Please note that any fractional degrees of freedom are due to Welch's test pooled degrees of freedom. The usage of Welch's test instead of Student's t-test is recommended (see http://daniellakens.blogspot.com/2015/01/always-usewelchs-t-test-instead-of.html).

correctly in the 'judgment after' condition, and 92% in the 'judgment before' condition, $t(114.79) = 0.75$, $p = .457$, $d = 0.14$; thus, the correct classification rates were roughly comparable between conditions. Participants misclassified on average 5 ('judgment after' condition) or 6 ('judgment before' condition) of 64 USs.

CS Ratings. Figure 4.6 presents participants' CS ratings averaged within US valence in the 'judgment after' and the 'judgment before' conditions. We analyzed these data the same way as in Experiment 1. The ANOVA showed the by now familiar valence main effect, $F(1,120) = 8.83$, $p = .004$, $\eta p^2 = 0.07$, replicating a typical EC effect. On average, participants rated CSs presented with positive USs as more likeable ($M = 0.31$, $SD = 0.72$) than CSs presented with negative USs ($M = 0.08$, $SD = 0.72$).

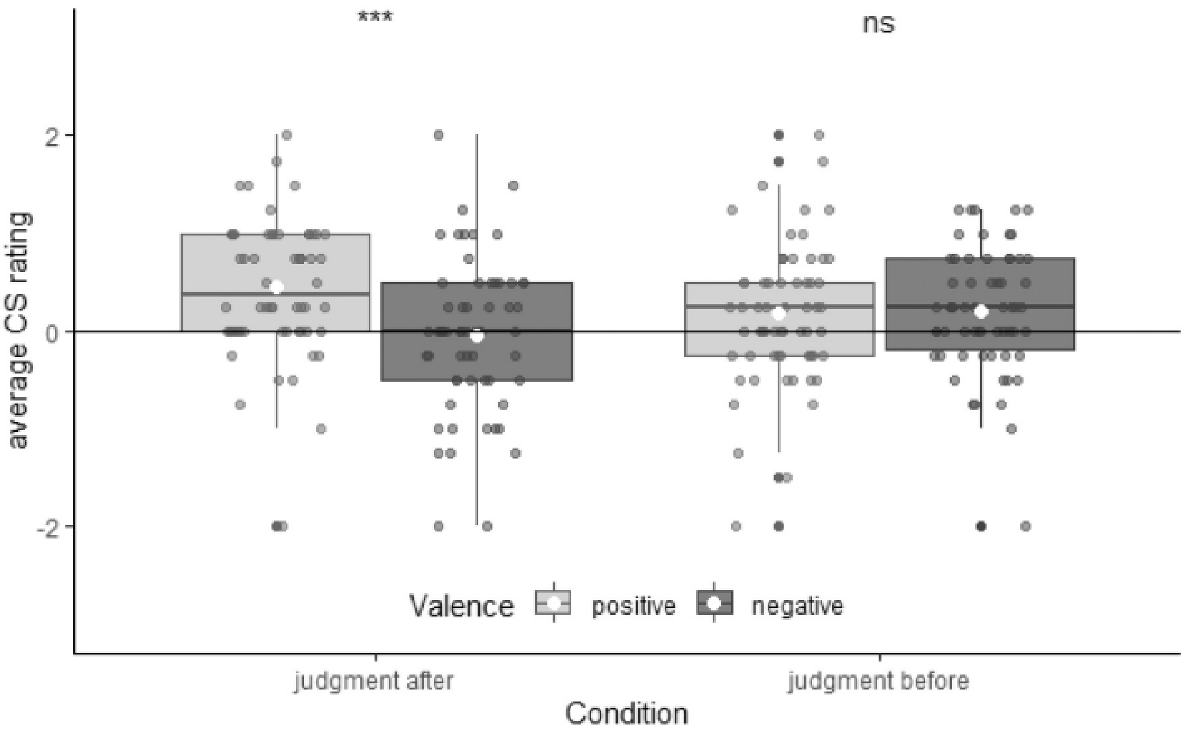
As Figure 4.6 suggests, this main effect was moderated by the judgment timing, $F(1, 120) = 11.09$, $p = .001$, $\eta p^2 = 0.08$. In the 'judgment after' condition, participants rated CSs presented with positive USs more positive ($M = 0.45$, $SD = 0.70$) than CSs presented with negative USs ($M = -0.04$, SD = 0.78), $t(59) = 3.75$, $M_{diff} = 0.50$, 95%CI [0.23, 0.76], $p < .001$, $d = 0.48$. This rating difference was even slightly reversed in the 'judgment before' condition, where the difference was no longer significant: Participants rated CSs presented with positive USs more negative ($M = 0.18$, $SD = 0.72$) than CSs presented with negative USs ($M = 0.21$, $SD = 0.63$), $t(61) = -0.32$, $M_{diff} = -0.03$, 95%CI [−0.20; 0.15], $p = .747$, $d = 0.04$. The main effect for timing condition was not significant, $F(1, 120) < 0.01$, $p = .949$, $\eta p^2 < 0.01$.

4.6.3 Bayesian Analysis

As Experiment 3 yielded a non-significant result using frequentist statistic methods, we again computed a Bayes factor analysis to provide further evidence for a null effect. Using the same program and procedure as in Experiments 1 and 2, the inverse Bayes factor analysis showed BF01 = 6.84 (i.e., moderate evidence for H0; Lee & Wagenmakers, 2013). Thus, our results are almost seven times more likely to be observed when assuming there is no effect than when assuming an effect (cf. Rouder et al., 2009).

Figure 4.6

Combined Box and Scatter plots of Participants' CS Ratings in Experiment 3 as a Function of US Valence (Within Participants) and Experimental Condition ('Judgment Before' vs. 'Judgment After'; Between Participants), Illustrating the Central Interaction.



Note. The simple effects for valence show that the difference between CS ratings in the 'judgment after' condition is highly significant but not significant in the 'judgment before' condition. Higher values indicate more positive evaluations. The box plots constitute the first quartile, third quartile, and median. The circle within a box plot indicates the mean. Each dot represents an evaluation. Observations outside the whiskers are outliers with >150% interquartile range distance from the next quartile.

4.6.4 Discussion

Experiment 3 critically extends the evidence for the present dual-force perspective on EC. Despite the spatio-temporal proximity of US and CS, no EC effect occurred when participants made a valence judgment in-between stimulus presentations. However, we observed a standard EC effect if participants made their judgment between trials and not between US-CS presentations. This pattern is difficult to explain by considering the judgment task as a mental load manipulation, and it is also difficult to explain the pattern from a how perspective, especially if one considers the almost demand-like nature of the judgment task. However, the pattern aligns with the assumption that the judgment forces the US and CS into separate episodes.

So far, we have only focused on top-down influences that force CS and US into separate mental episodes. This focus results from EC's roots in a behavioristic tradition manipulating environmental regularities (Hughes, Mattavelli et al., 2018). Our next experiment shows that we can re-establish the EC effect by additional top-down input.

4.7 Experiment 4: Top-Down Establishment of Mental Episodes

Experiment 4 aimed to show a standard EC effect in the 'judgment before' condition by re-integrating CS and US through top-down input on the organism's side. To manipulate the respective top-down force, we 4 directly manipulated the mental episodes using instructions.

4.7.1 Method

Participants. As pre-registered, we collected data from 200 participants on a University campus who participated for payment of 2 Euro (about $2.30). We increased the sample size as the manipulation by instructions may have less impact than the procedural manipulation (i.e., changing the judgment order on each trial). As pre-registered, we excluded one participant who did not finish the study. We excluded 13 further participants as they stated low concentration.[13] We excluded one participant who indicated to speak/understand Japanese/Chinese. Finally, we excluded one more participant because the person was younger than 18 years and could not provide informed consent for participation. Therefore, we retained 184 participants for analysis (142 women, 35 men, 1 other, 6 prefer not to say; M_{age} = 22.41, SD_{age} = 3.72). A mixed ANOVA with 91 participants in the 'pairing instruction' condition and 93 participants in the 'judgment instruction' condition (N = 184) would be

[13] Instead of asking participants if we could use their data (as stated in the preregistration), they were asked about their level of concentration. As in the three preceding experiments, participants who stated low, very low or extremely low concentration were excluded.

sensitive to effects of $\eta p^2 = 0.04$ with 80% power ($\alpha = 0.05$, $r = 0.50$). This means the experiment would not be able to reliably detect effects smaller than $\eta p^2 = 0.04$.

Design. The participants were assigned to one of two groups of a 2 (instruction: judgment vs. pairing; between) by 2 (US valence: positive vs. negative; within) mixed design.

Procedure. The materials were the same as in Experiments 1 to 3. The procedure was also highly similar with three modifications. First, participants provided a judgment in-between CS and US in both conditions; that is, we replicated Experiment 3's 'judgment before' condition. Second, critically, we used different instructions to force US and CS into the same mental episode despite the judgment separating them (i.e., 'pairing condition') or to force US and CS into different mental episodes (i.e., 'judgment condition'); and third, every US-CS-presentation trial was preceded by the information "next pairing" ('pairing instruction' condition) or 'next judgment' ('judgment instruction' condition). To further strengthen the pairing in the 'pairing instruction' condition, participants read 'end of pairing' after each trial (1000 ms), while participants in the 'judgment instruction' were presented a black blank (1000 ms). The computer program assigned participants alternatingly to one of two between-conditions.

In the 'pairing instruction' condition, the instructions informed participants to pay close attention to the pairing of US and CS: 'In the second part of the study,[14] you will be shown pairs of pictures. First, you will see the picture of an animal and, after a choice task, the related Chinese Ideograph. Your task is to decide whether the first picture of the pairing (the picture of the animal) is negative or positive. The picture pairings are repeatedly shown during the study. Please observe both pictures of the pairing attentively.'

In the 'judgment instruction' condition, the instructions informed participants to focus on the US while ignoring the subsequent CS: 'In the second part of the study, you will be shown single pictures. The picture of an animal is always followed by a choice task. It is your task to decide

[14] The first part refers to the preratings of the USs.

whether the picture of the animal is negative or positive. The single pictures are repeatedly shown during the study. The Chinese Ideographs can be ignored for this study.'

Both conditions used the same backward conditioning procedure introduced in Experiments 1-3.

4.7.2 Results

Valence Judgments. In the 'judgment instruction' condition, participants, on average, pressed a key within 825 ms (*SD* = 229 ms). In the 'pairings instruction' condition, participants, on average, pressed a key within 876 ms (*SD* = 372 ms). These response latencies did not differ significantly, $t(149.06) = -1.11$, $p = .269$, $d = 0.16$. In addition, participants classified 96% of the US correctly in the "judgment instruction" condition, and 97% in the "pairing instruction" condition, $t(178.72) = -0.47$, $p = .640$, d = 0.07. On average, participants in both conditions misclassified 2 USs.

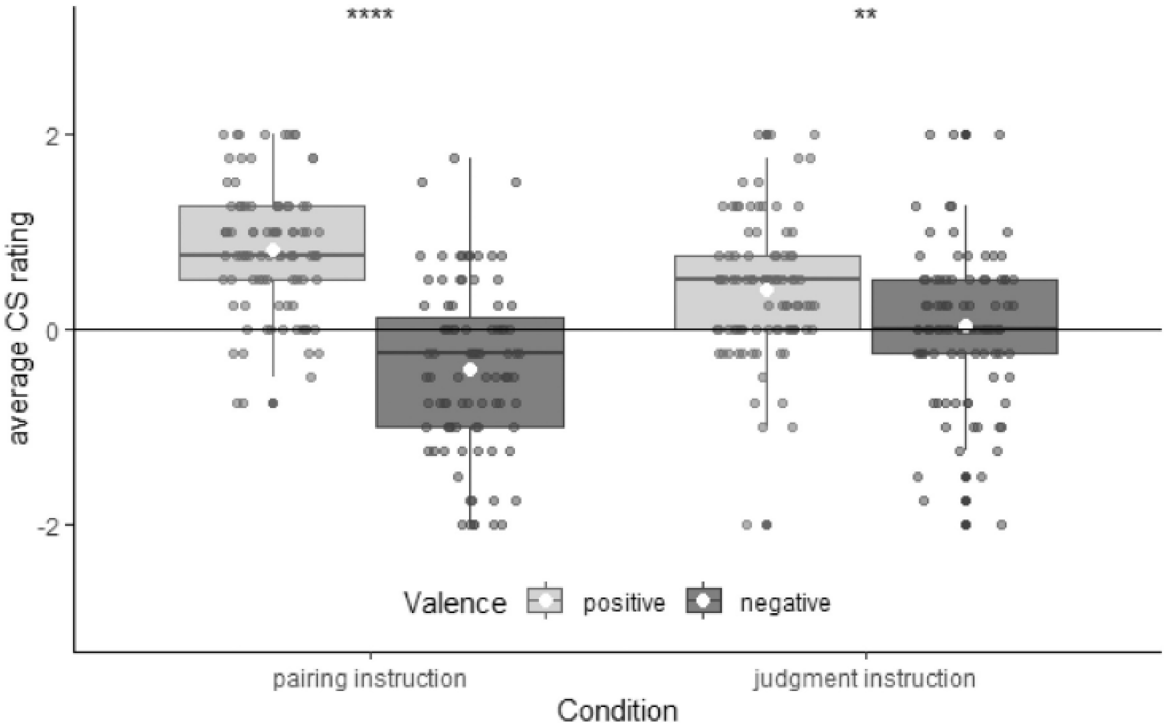
CS Ratings. Figure 4.7 presents participants' CS ratings averaged within US valence in the 'pairing instruction' and the 'judgment instruction' conditions. We analyzed these data with an instruction ("pairing" vs. "judgment") by US ("positive" vs. "negative") mixed ANOVA with repeated measures on the latter factor.

The ANOVA showed a valence main effect, $F(1, 182) = 101.36$, $p < .001$, $\eta p^2 = 0.36$. On average, participants rated CSs presented with positive USs as more likeable ($M = 0.61$, $SD = 0.70$) than CSs presented with negative USs ($M = -0.18$, $SD = 0.83$). As Figure 4.7 suggests, this main effect was moderated by the instruction conditions, $F(1, 182) = 28.39$, $p < .001$, $\eta p^2 = 0.13$.

In the 'pairing instruction' condition, participants rated CSs presented with positive USs more positive ($M = 0.82$, $SD = 0.66$) than CSs presented with negative USs ($M = -0.40$, $SD = 0.86$), $t(90) = 10.15$, $M_{diff} = 1.22$, 95%CI [0.98, 1.46], $p < .001$, $d = 1.06$. As the interaction indicates, this difference was significantly smaller in the 'judgment instruction' condition.

Figure 4.7

Combined Box and Scatter plots of Participants' CS Ratings in Experiment 4 as a Function of US Valence (Within Participants) and Experimental Condition ('Pairing Instruction' vs. 'Judgment Instruction'; Between Participants), Illustrating the Central Interaction.



Note. The simple effects for valence show that the difference between CS ratings in the 'pairing judgment' instruction condition is highly significant. The difference in the 'instruction judgment' condition is smaller but still significant. Higher values indicate more positive evaluations. The box plots constitute the first quartile, third quartile, and median. The circle within a box plot indicates the mean. Each dot represents an evaluation. Observations outside the whiskers are outliers with >150% interquartile range distance from the next quartile.

Different from the previous three experiments, the difference was still significant: Participants rated CSs presented with positive USs more positive (*M* = 0.41, *SD* = 0.69) than CSs presented with negative USs (*M* = 0.03, *SD* = 0.75), $t(92) = 3.62$, $M_{diff} = 0.38$, 95%CI [0.17; 0.58], $p < .001$, $d = 0.38$. The ANOVA main effect for condition was not significant, $F(1, 182) = 0.02$, $p = .892$, $\eta p^2 < 0.01$.

4.7.3 Discussion

Experiment 4 shows that the top-down instructions to perceive US and CS as a pairing re-establishes the EC effect despite the valence judgment in-between US and CS; thus, the top-down input intended to force US and CS into the same mental episode led to a strong EC effect. Experiment 4 thereby provides evidence that it is not the judgment between the US and CS *per se*; if the top-down influence forces US and CS into the same episode, the disruption of the Gestalt psychological influences (e.g., #US – judgment – CS') has no influence. Together with Experiments 1 to 3, this provides evidence for the interplay of bottom-up and top-down processes in establishing the necessary conditions for EC effects.

Different from Experiments 1 to 3, we also found a significant EC effect in the judgment instruction condition. In hindsight, this difference from the previous experiments likely follows from a procedural change in Experiment 4. To reaffirm the pairing instruction on each trial, we included a 'next pairing' screen in-between each US-CS trial. We also included a 'next judgment' screen in the judgment instruction condition to keep the conditions comparable. From a Gestalt perspective, this establishes a temporal pattern of ‚US – judgment – CS – pause'; similarly, the 'next judgment' screen may serve as a clear marker for a given mental episode. Post-hoc, we consider this a likely explanation within our theoretical framework for the variations between Experiment 4 and 1 to 3. In addition, Experiment 4 had the highest statistical power to detect an EC effect.

Thereby, the EC effect tested by the simple effects may represent a mixture of the pairing instruction (i.e., top-down) and the experiments' configural setup (i.e., bottom-up). However, the central argument of the present dual-force perspective, here and in Experiments 1-3, is the interaction, not the differential simple effects. This interaction follows from the differential impact of the 'next pairing' vs. 'next judgment' instruction on the EC effect, clearly showing the top-down component within the trial setup's configural constituents. One may argue that the present manipulation is not very subtle; it constitutes an explicit manipulation of mental episode formation. However, one may construe more refined manipulations, for example, world knowledge (e.g., red wine goes with meat and white wine goes with fish), schemas (e.g., foods and restaurants vs. foods and park benches), or stereotypes (e.g., cherry blossoms and kanjis, as both stereotypically Japanese stimuli) could serve as more subtle exemplars of top-down influences that force CS and US into the same or separate mental episodes. However, as world knowledge and schemas may differ between participants, we decided to use this unambiguous manipulation by instruction of a top-down force.

Additionally, in the ‚judgment instruction' condition we asked participants to ignore the CS. This change in instructions may have prevented participants from processing the CS. However, experiments 1-3 already show that segmenting stimuli into different episodes is possible without instructing them to ignore the CS. The important finding here is that the pairing instruction re-establishes the mental episode.

Thus, with Experiments 1 to 3, we provided evidence that top-down input can force CS and US into different mental episodes. Experiment 4 shows that other top-down inputs can force CS and US into the same mental episode while keeping the bottom-up forces constant. However, three remaining points need to be addressed. First, our manipulation may be especially sensitive to the backward conditioning procedure we used in Experiments 1-4. Second, the question in-between the presentation of CS and US already forces the stimuli into separate episodes, rendering top-down input redundant. Third, the instructions of the former experiments may leave an attentional confound because the focus in the “no judgment” conditions was laid on CS and US, while the focus in the judgment conditions was on the US alone. Thus, we conducted two additional experiments to address these pending points.

4.8 Experiment 5: Forward Conditioning

So far, we have provided evidence for the suggested dual-force perspective using backward conditioning procedures to allow US judgments. Experiment 5 aimed to show that the proposed dual-force perspective also applies to forward conditioning. Therefore, we replicated Experiment 1 using the more typical CS – US sequence.

4.8.1 Pretest

Due to the COVID-19 pandemic, Experiment 5 was conducted online. To ensure that EC can also be found in an online setting, we conducted a pretest with 75 participants (sequential analysis; we pre-registered 150 participants for $\alpha > 0.294$). Two participants were excluded due to speaking or understanding Japanese/Chinese. We analyzed the data of the 73 remaining participants (55 women, 18 men; $M_{age} = 40.65$, $SD_{age} = 12.26$) using a paired t-test. We found a significant difference between CSs presented with positive USs ($M = 0.64$, $SD = 0.68$) and CSs presented with negative USs (M = 0.35, $SD = 0.77$), $t(72) = 2.50$, $M_{diff} = 0.30$, 95%CI [0.06, 0.54], $p = .015$, $d = 0.42$. Thus, the pretest data showed that we can detect EC effects in this online setting.

4.8.2 Method

Participants. As pre-registered, we collected data from 150 participants (after exclusion[15]) on Prolific who participated for payment of ₤1.63 (about $2.21). As pre-registered, we excluded one participant who requested not to use their data. We excluded one further participant who indicated to speak/understand Japanese/Chinese. Therefore, we retained 148 participants for analysis (111 women, 35 men, 1 diverse, 1 non-binary; $M_{age} = 39.75$, $SD_{age} = 15.36$). A mixed ANOVA with 75 participants in the 'location judgment' condition and 73 participants in the 'no judgment' condition (N = 148) would be sensitive to effects of $\eta p^2 = 0.05$ with 80% power ($\alpha = 0.05$, $r = 0.50$). This means the experiment would not be able to reliably detect effects smaller than $\eta p^2 = 0.05$.

Design. The participants were assigned to one of two groups of a 2 (instruction: location judgment vs. no judgment; between) by 2 (US valence: positive vs. negative; within) mixed design.

Procedure. The materials were the same as in Experiments 1 to 4. The procedure was identical to Experiment 1 except for the forward conditioning modification (i.e., presenting the CS before the US; cf. Figure 4.2). Participants in the 'location judgment' condition thus had to judge the

[15] Prolific automatically continues collecting data when participants did not finish the study. Thus, out of 168 participants, data from 156 participants was collected (no data is available for 12 participants that returned the study). 6 of these remaining participants did not finish the study and were excluded by Prolific, leaving us with the preregistered 150 participants.

location of the CS and not the US, as in Experiment 1. Additionally, participants did not choose the presented USs idiosyncratically: Due to the change to an online setting, we aimed to keep the experiment's duration short. The choice of USs was based on two steps: First, we analyzed the US ratings from Experiments 1-4 and selected the 4 USs with the most negative ratings across experiments and the 4 USs with the most positive ratings across experiments. Second, as the animal pictures used IAPS images that are not permitted with open-source survey use (like Prolific; see https://csea.phhp.ufl.edu/Media.html) for online research, we compared those USs with pictures of the Open Affective Standardized Image Set (OASIS; Kurdi et al., 2017) and tried to find pictures that were comparable with the IAPS pictures in appearance and valence. When we did not find a picture comparable in appearance and valence, we used a picture that was at least comparable in valence. The set of negative stimuli retrieved from OASIS contained Cockroach 1, Snake 5, Spider 2, and Wolf 2; the set of positive stimuli consisted of Cat 5, Dog 6, Dog 12, and Penguins 2. The computer program assigned participants alternatingly to one of two between-conditions.

In the 'no judgment' condition, the instructions informed participants to pay close attention to the presented stimuli: 'You will see pictures of Japanese characters and animals. Your task is to watch the presented pictures attentively.' In the 'location judgment' condition, in addition to the instruction to focus on the stimuli, participants were informed to make a judgment about the CS: 'You will see pictures of Japanese characters and animals. Your task is to watch the presented pictures attentively and decide whether the Japanese character's picture was presented on the top or the bottom half of the screen using the 'A' key for top and the 'L' key for bottom.' Thus, we asked participants to pay attention to CS and US to avoid an attentional confound between conditions.

Finally, participants needed to press the space bar after a CS – US or a CS – judgment – US sequence to ensure that they observed the stimulus presentation.

4.8.3 Results

Location Judgments. In the 'no judgment' condition, the presentation time between CS and US was set to 990 ms (based on the third additional Experiment, see Supplementary Material). In the 'location judgment' condition, participants, on average, pressed a key within 963 ms (*SD* = 297 ms). These response latencies did not differ significantly from the 'no judgment' condition, $t(74) = 0.79$, $p = .433$, $d = 0.09$. In addition, participants classified 99% of the CSs correctly in the 'location judgment' condition. On average, participants misclassified <1 CS.

CS Ratings. Figure 4.8 presents participants' CS ratings averaged within US valence in the 'no judgment' and the 'location judgment' conditions. We analyzed these data with a judgment (no judgment vs. location judgment) by US (positive vs. negative) mixed ANOVA with repeated measures on the latter factor.

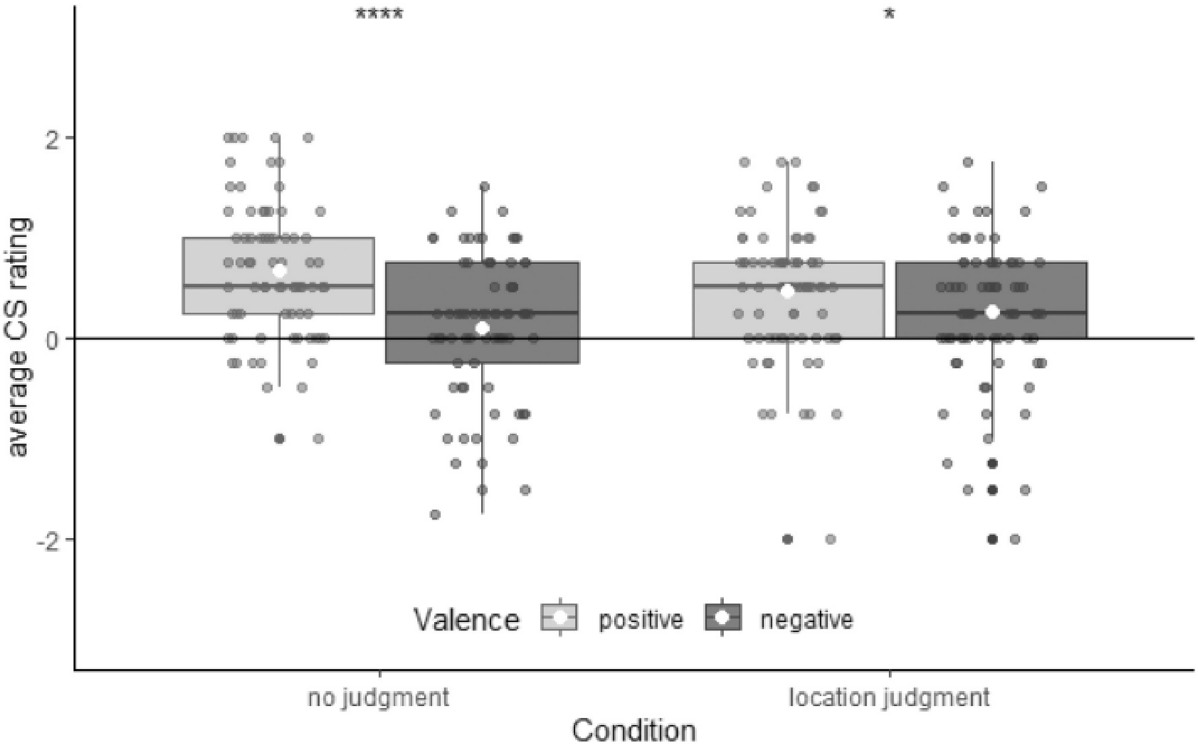
The ANOVA showed a valence main effect, $F(1, 146) = 26.81$, $p < .001$, $\eta p^2 = 0.16$. On average, participants rated CSs presented with positive USs as more likeable ($M = 0.57$, $SD = 0.68$) than CSs presented with negative USs ($M = 0.18$, $SD = 0.72$).

As Figure 4.8 suggests, this main effect was moderated by the judgment conditions, $F(1, 146) = 5.65$, $p = .019$, $\eta p^2 = 0.04$. Participants in the "no judgment" condition rated CSs presented with positive USs more positive ($M = 0.68$, $SD = 0.68$) than CSs presented with negative USs ($M = 0.10$, $SD = 0.73$), t(72) = 4.72, $M_{diff} = 0.58$, 95%CI [0.33, 0.82], $p < .001$, $d = 0.81$. As the interaction indicates, this difference was significantly smaller in the "location judgment" condition.

As in Experiment 4, this difference was still significant: Participants rated CSs presented with positive USs more positive ($M = 0.47$, $SD = 0.67$) than CSs presented with negative USs ($M = 0.25$, $SD = 0.71$), $t(74) = 2.32$, $M_{diff} = 0.21$, 95%CI [0.03, 0.40], $p < .001$, $d = 0.31$. The ANOVA main effect for condition was not significant, $F(1, 146) = 0.13$, $p = .723$, $\eta p^2 < 0.01$.

Figure 4.8

Combined Box and Scatter plots of Participants' CS Ratings in Experiment 5 as a Function of US Valence (Within Participants) and Experimental Condition ('No Judgment' vs. 'Location Judgment'; Between Participants), Illustrating the Central Interaction.



Note. The simple effects for valence show that the difference between CS ratings in the 'no judgment' instruction condition is highly significant. The difference in the 'location judgment' condition is significantly smaller but still significant. Higher values indicate more positive evaluations. The box plots constitute the first quartile, third quartile, and median. The circle within a box plot indicates the mean. Each dot represents an evaluation. Observations outside the whiskers are outliers with >150% interquartile range distance from the next quartile.

4.8.4 Discussion

Experiment 5 replicates the results of Experiment 1 using a forward conditioning procedure. Judging the location of the CS (bottom vs. top of the screen) has similar effects compared to the location judgment of the US. Thus, the order of stimulus presentation does not influence the segmentation of CS and US.

Similar to Experiment 4, we also found a significant EC effect in the judgment condition (cf., judgment instruction condition in Experiment 4). As in Experiment 4, this difference to Experiment 1-3 may follow from a procedural change in Experiment 5. While in Experiments 1 to 3, the ITI was set to 2500 ms, the online setting of Experiment 5 required inserting a spacebar press after each CS-US presentation to increase the likelihood that participants did not leave their computers while performing the study. As we did not know how fast participants would press the spacebar, participants could not proceed before 2500 ms, as in Experiments 1 to 3 (i.e., 2500 ms). Adding the spacebar press to the fixed ITI increased the ITI to, on average, 4105 ms. Again, from a Gestalt perspective, this establishes a temporal pattern of 'US – judgment – CS – pause', thereby establishing a stronger bottom-up force across conditions than in Experiments 1 to 3. However, again, the central argument for our dual-force perspective is the interaction, not the differential simple effects.

Finally, the overall EC effect in Experiment 5 is weaker than in Experiments 1-4. One straightforward explanation is that we conducted the experiment online. In particular, the problem of inattention, for example, due to media distraction (Chandler et al., 2014) or multitasking (Necka et al., 2016), can negatively affect the quality of data (Newman et al., 2021).

4.9 Experiment 6: Comparable Bottom-Up Influence

The previous experiments showed that a judgment between CS and US significantly reduces or eliminates the EC effect. We argued that this top-down influence (i.e., the judgment) forces CS and US into separate mental episodes (Fiedler et al., 2005). However, one might argue that presenting the question prompt in-between stimuli is a bottom-up influence that forces the separate mental episodes, while the judgment per se would be coincidental and potentially unnecessary. This argument would only leave Experiment 4 as evidence for the dual-force perspective.

To investigate whether the mere bottom-up input (i.e., the question prompt) forces CS and US into distinct episodes, we replicated Experiment 5 but included a comparable bottom-up input to the question prompt in the 'no judgment' condition. Concretely, we changed the white blank between CS and US in the 'no judgment' condition to a blank containing grey bars of the same length and color intensity as in the judgment condition. In the 'location judgment' condition, the question prompt (using grey letters) was presented in-between stimulus presentation. This way, we kept the bottom-up input constant between conditions, as reading, processing, and reacting are on the organism's side.

4.9.1 Method

Participants. Due to the COVID-19 pandemic, this experiment was also conducted online. As preregistered, we collected data from 300 participants (after exclusion[16]) on Prolific who participated for payment of ₤1.67 (about $2.27). As preregistered, we excluded eight participants who indicated to speak/understand Japanese/Chinese. Therefore, we retained 292 participants for analysis (226 women, 65 men, 1 prefer not to say; M_{age} = 40.57, SD_{age} = 14.11). A mixed ANOVA with 144 participants in the 'location judgment' condition and 148 participants in the 'no judgment' condition (N = 292) would be sensitive to effects of $\eta p^2 = 0.03$ with 80% power ($\alpha = 0.05$, $r = 0.50$). This means the experiment would not be able to reliably detect effects smaller than $\eta p^2 = 0.03$.

Design. The participants were assigned to one of two groups of a 2 (judgment: location judgment vs. no judgment; between) by 2 (US valence: positive vs. negative; within) mixed design.

Procedure. The materials were the same as in Experiment 5. The procedure was highly similar to Experiment 5 except for modifying the blank in-between CS and US in the 'no judgment' condition. Participants in the 'location judgment' condition again judged the CS location. In Experiment 6, the letters of the judgment question were presented in light grey instead of black to keep the input

[16] Data collection of 338 participants due to incomplete data sets on Prolific: 29 participants returned the study, 9 participants aborted the study early, and were excluded by Prolific (i.e., Prolific does not count unfinished data as collected data; as preregistered, we excluded these data sets). Thus, complete data sets of 300 participants were collected.

between conditions constant. We decided to use light grey (hex code: #D3D3D3) instead of black color (hex code: #000000) to investigate whether the top-down input segments CS and US. Black bars, however, may constitute a relatively strong bottom-up influence. Using light grey color provides thus the most straightforward way to test top-down influences while keeping the bottom-up input constant.

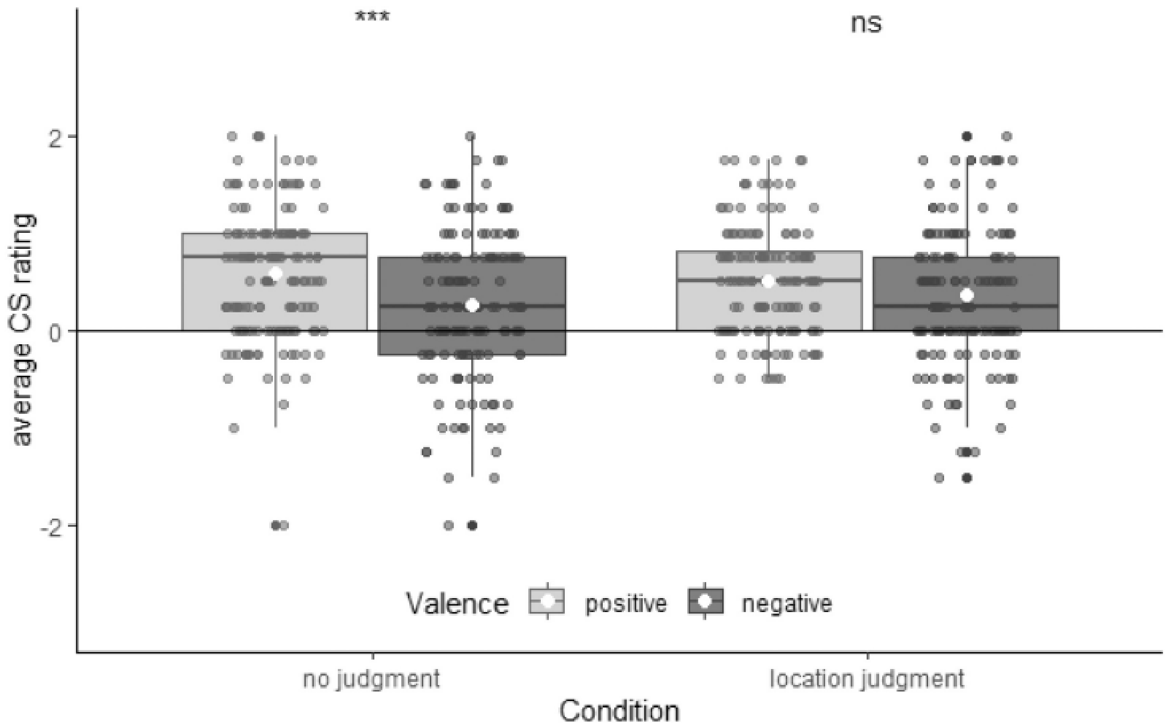
4.9.2 Results

Location Judgments. As in Experiment 5, the time between CS and US presentation was 990 ms in the 'no judgment' condition. In the 'location judgment' condition, participants, on average, pressed a key within 1088 ms (*SD* = 744 ms). These response latencies were comparable to the specified interval in the 'no judgment' condition, $t(143) = 1.59$, $p = .113$, $d = 0.13$. In addition, participants classified 99% of the CSs correctly in the 'location judgment' condition (i.e., participants misclassified <1 CS on average).

CS Ratings. Figure 4.9 presents participants' CS ratings averaged within US valence in the 'no judgment' and the 'location judgment' conditions. We analyzed these data with a judgment ('no judgment' vs. 'location judgment') by US ('positive' vs. 'negative') mixed ANOVA with repeated measures on the latter factor. The ANOVA showed a valence main effect, $F(1, 290) = 18.51$, $p < .001$, $\eta p^2 = 0.06$. On average, participants rated CSs presented with positive USs as more likeable ($M = 0.55$, $SD = 0.63$) than CSs presented with negative USs ($M = 0.31$, $SD = 0.74$). As Figure 4.9 suggests, this main effect was moderated by the judgment conditions. However, the interaction was not significant, $F(1, 290) = 2.60$, $p = .108$, $\eta p^2 = 0.01$. This is the sole case when we need to fall back to the preregistered nested contrast analysis, as the ANOVA might not have enough power to detect the hypothesized attenuate. In the 'no judgment' condition, participants rated CSs presented with positive USs more positive ($M = 0.58$, $SD = 0.67$) than CSs presented with negative USs ($M = 0.26$, $SD = 0.76$), $\beta = 0.32$, 95%CI [0.17, 0.48], $p < .001$. This difference was not significant in the 'location judgment' condition. Here, participants rated CSs presented with positive USs less positive ($M = 0.51$, $SD = 0.58$) and less negative ($M = 0.36$, $SD = 0.71$) than in the 'no judgment' condition, $\beta = 0.15$, 95%CI [− 0.01, 0.30], *p* = .071. Further information regarding the nested contrast analysis can be found in the supplementary material. The ANOVA main effect for condition was not significant, $F(1, 290) = 0.07$, $p = .796$, $\eta p^2 < 0.01$.

Figure 4.9

Combined Box and Scatter plots of Participants' CS Ratings in Experiment 6 as a Function of US Valence (Within Participants) and Experimental Condition ('No Judgment' vs. 'Location Judgment'; Between Participants), Illustrating the Rating Difference Between Conditions.



Note. The simple effects for valence show that the difference between CS ratings in the 'no judgment' instruction condition is highly significant. The difference in the 'location judgment' condition is not significant. Higher values indicate more positive evaluations. The box plots constitute the first quartile, third quartile, and median. The circle within a box plot indicates the mean. Each dot represents an evaluation. Observations outside the whiskers are outliers with >150% interquartile range distance from the next quartile.

4.9.3 Bayesian Analysis

As Experiment 6 yielded a non-significant result using frequentist statistic methods, we again computed a Bayes factor analysis to provide further evidence for a null effect. Using the same program and procedure as in Experiments 1-3, the inverse Bayes factor analysis showed BF01 = 1.23 (i.e., anecdotal evidence for H0; Lee & Wagenmakers, 2013). Thus, our results are slightly more likely to be observed when assuming there is no effect than when assuming an effect (cf. Rouder et al., 2009).

4.9.4 Discussion

One concern of Experiment 1-5 was that the bottom-up input in the conditions where participants had to make a judgment is the relevant influence, leading to reduced or eliminated EC effects. While this concern does not question our answer to the *when* question per se, it would weaken the contribution of the dual-force perspective. To address this concern, Experiment 6 removed the bottom-up confound potentially present in Experiments 1 to 5. Given the strong EC effect in the no judgment condition despite the presentation of a 'block' (i.e., grey bars) between CS and US, the reduced EC effects in Experiments 1-5' judgment conditions were not solely based on the question text between the stimulus presentation (i.e., the bottom-up input). Instead, the reduction also depended on the input from the organism's side. By contrast, the judgment reduced the EC effect to non-significance, which was our pre-registered pattern (cf. Schad et al., 2020).

4.10 General Discussion

We started with the observation that EC research has provided many insights into the processes by which people acquire their likes and dislikes due to the co-occurrence of two stimuli; it addressed the how question (see De Houwer et al., 2020; Jones et al., 2010). However, in comparison, the question *when* people learn from the co-occurrence of two stimuli received less attention. From a behavioristic perspective, the *when* question is answered by spatio-temporal contiguity. EC should occur if CS and US co-occur (repeatedly) in space and time. We argued that this behavioristic perspective, with its focus on environmental regularities, needs to be enriched by a social-cognitive perspective to answer the *when* question. We suggested the construct of mental episodes (Fiedler et al., 2005) on the organism's side as the necessary condition for EC, which is highly similar and potential exchangeable with other constructs in cognitive psychology (e.g., event models, Radvansky, 2012; event files, Hommel, 2004).

To delineate determinants of a mental episode, we presented the dual-force model by Fiedler, 2001; Fiedler & Bless, 2000. Accordingly, accommodative forces (i.e., environmental, bottom-up influences) and assimilative forces (i.e., organismic, top-down influences) influence the formation of mental episodes. The most decisive bottom-up influence is probably two stimuli's repeated co-occurrence in space and time. However, from the present dual-force perspective, we predicted and showed empirically that a top-down segmentation of US and CS forces stimuli into separate mental episodes, despite the strong bottom-up influence. Concretely, building on the work by Fiedler et al. (2005) and Radvansky (2012), we hypothesized that a judgment in-between US and CS would force CS and US into different episodes.

This prediction distinguishes the present *when* question from the *how* question because prototypical models of the underlying processes cannot explain the absence of EC effects in these cases without substantial additional assumptions. Across six experiments, we kept the spatial and temporal proximity of US and CS constant. Participants observed USs (animal pictures) followed by CSs (kanjis). In Experiment 1, participants judged the US's screen location (top vs. bottom) before the CS presentation. This location judgment significantly reduced the EC effect, which was not significant. In Experiment 2, participants judged the US valence instead of US location and thereby ruled out that a lack of participants' encoding of US valence leads to the lack of an EC effect. In addition, from a propositional perspective, a strong EC effect should follow when US valence is actively confirmed. Nevertheless, the judgment significantly reduced the EC effect to non-significance. Experiment 3 showed that it is not the judgment per se but the timing of the judgment in-between the US and CS. Participants provided valence judgments in all conditions, but participants in one condition made judgments between the US-CS presentation and in the other condition after the US-CS presentation.

We predicted and found an EC effect in the latter condition, but the effect was again reduced to non-significance when participants judged the US valence in-between US and CS. Experiment 4 instructed participants to focus on the US-CS pairing or the US judgment, showing that top-down forces can both prevent and foster CS and US encoding into the same mental episode. The pairing instruction condition showed a strong EC effect despite the valence judgment in-between US and CS, which was significantly reduced in the judgment construction condition. Experiment 5 replicated Experiment 1 using a forward conditioning procedure. Finally, Experiment 6 addressed a concern about whether it is indeed the top-down input and not the bottom-up input of the question prompt that segments CS and US into a stream of events, that is, into distinct episodes; without providing a judgment, we found a strong EC, even when CS and US were separated by a set of bars the same size as the question prompt. When participants provided a judgment, this effect was again reduced to non-significance. Together, these six experiments support a dual-force framework to answer the *when* question, based on the interplay of bottom-up and top-down forces, determining whether CS and US are encoded into the same or separate mental episodes.

We thereby suggest a new perspective, and this requires a close inspection if the present results may not follow from established principles in EC research. Experiments 2 and 3 already ruled out the lack of a valence focus (Gast & Rothermund, 2011) and mental load (Stahl & Aust, 2018). Nevertheless, one might argue that the judgment setup creates demand effects. However, if anything, judging US valence (Experiments 2-4) should reinforce the demand that the CS should be seen as positive (or negative) after a ‘positive’ (or ‘negative’) judgment. In addition, all experiments manipulated the critical variables between participants, which makes inferring the intentions of the experimental manipulations difficult. Another alternative is that the present results follow from a social comparison contrast effect (e.g., Mussweiler, 2003; Suls et al., 2002). Judging US valence might create a situation where the CS appears different from the US, and thereby less likable after a positive US and more likable after a negative US (cf. Unkelbach & Fiedler, 2016, for a direct test of this prediction). However, this argument cannot explain the location judgment effect in Experiment 1 and the effect of the pairing instruction in Experiment 4. Beyond these alternative explanations, the greatest challenge is whether our data might not follow from answers to the *how* questions.

In the remainder, we want to discuss the relation between the *when* and the *how* questions, connect the present empirical evidence with existing work, and highlight again why the social-cognitive perspective on the *when* question, that is, beyond the behavioristic spatio-temporal contiguity, is helpful to understand the EC phenomenon.

4.10.1 The Relation of the When and the How Question

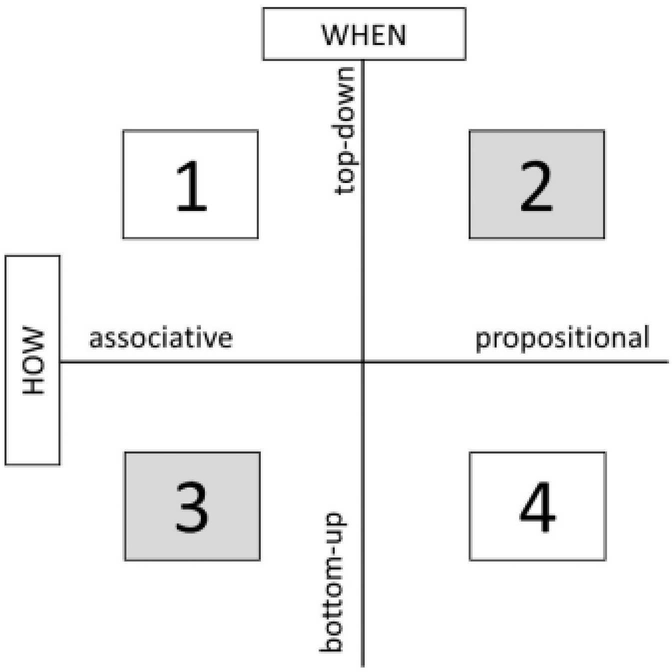
In the introduction, we argued that the *when* and the *how* questions are independent. If CS and US are encoded into the same mental episode, EC effects might occur due to associative processes, propositional processes, misattribution, memory, and so forth. As long as this encoding is due to bottom-up influences, it seems easy to accept that the co-occurrence could be represented, for example, as an 'association' or enriched to a 'proposition'. The other way around is more difficult to accept. If the encoding or lack thereof is due to top-down influences, it seems face-valid that the proper EC process must be propositional. In other words, one might argue that the present investigation represents a variant of propositional learning experiments. Let us scrutinize this argument for explanations based on propositions and associations.

Specifically, there is good evidence that propositional and thereby 'top-down input' can diminish or even reverse EC effects, for example, by providing a relational qualifier in-between CS and US (e.g., Förderer & Unkelbach, 2012; Högden & Unkelbach, 2021; Hughes, Ye, Van Dessel et al., 2019; Moran & Bar-Anan, 2013). However, upon closer inspection, the data from Experiments 2 to 4 do not follow from a propositional learning account. It would be difficult to explain why evaluating US valence, which reinforces US valence, should diminish EC effects (Experiment 2) and why the timing of the judgment should matter (Experiment 3). Alternatively, one may argue that what we call the top-down force is another label for 'associations' (i.e., the unqualified link between CS and US). Thus, the judgment manipulation might not force US and CS into separate mental episodes but simply prevent CS-US 'associations' (Gawronski & Bodenhausen, 2014, 2018). However, in Experiment 4, we observed EC effects when we asked participants to perceive US and CS as a pairing. One would need to argue that the judgment in-between US and CS prevents associations, but Experiment 4's EC effects represent propositional EC effects. In other words, every single experiment might be explained by assuming either propositional or associative processes; yet, given the experiments' high similarity, it is then unclear why Experiments 1 to 3 should rely on the disruption of associative processes but do not follow propositional processes, and Experiment 4 then leads to a propositional processes. Thus, we believe the pattern across experiments is difficult to explain from the dominant answers to the how question without substantial additional assumptions.

To illustrate this point further, Figure 4.10 shows the four potential combinations of the *when* and the *how* question.

Figure 4.10

Illustrating the Four Possible Combinations of the When and the How Question.



Note. For the purpose of the graph, the how question is highly simplified. There are several other answers to the how question (see Balas & Bar-Anan, 2018).

If our answer to the *when* question would be redundant with the *how* question, only quadrants #2 and #3 (or #1 and #4) should exist empirically. However, as stated above, both propositional and associative processes resulting from bottom-up input are frequently discussed in the literature. In other words, quadrants #2, #3, and #4 seem to be accepted. The remaining quadrant is #1; it seems challenging to find examples of purely top-down driven mental episodes that rely on associative processes to explain EC effects. However, this challenge is not a theoretical impossibility. For example, there is good evidence for EC effects via instructions; that is, EC without bottom-up input. Nevertheless, one might argue that the instruction places CS and US into the same mental episode, and the process that leads to the EC effect within the episode could be associative. We hasten to add that we believe a single process approach to EC has many advantages (see Hu et al., 2017a; Hu et al., 2017b), as it would also simplify our answer to the *when* question. However, at the moment, we consider the questions to be orthogonal. We thus believe that the present results are difficult to explain from both an associative and a propositional perspective (i.e., the most prominent answers to the *how* question). By a priori considering the *when* question, though, the present experiments' data pattern is directly predicted (see pre-registrations).

4.10.2 Connections with Existing Work

The differentiation between bottom-up and top-down input on EC is not new (cf. Fiedler & Unkelbach, 2011). However, applying a dual-process model to mental episodes in EC provides a framework for research regarding the *when* question. We will distinguish the organism's input on EC from the organism's input on mental representations by reviewing some of the literature examining the role of memory, attention, and awareness on EC, providing a *when* perspective of the findings.

Awareness. Previous research on EC assumes that regularities in the environment determine the underlying mental processes (Jones et al., 2010; Sweldens et al., 2010). For example, Hütter and Sweldens (2013) argued that simultaneous stimulus presentation leads to a misattribution process and EC effects of which the organism is 'unaware'. Conversely, consecutive stimulus presentation should lead only to EC effects under conditions of accessible memory representation, thereby specifying a propositional process. In our terminology, the simultaneous presentations represent a stronger bottom-up input, and thereby, less input from the organism would be needed.

Attention. Another example of how top-down input may influence mental episodes is Corneille et al. (2009). The authors manipulated the processing goal of participants using a task that either primed a similarity-focus or a difference-focus. They found that participants primed with a similarity-focus showed a stronger EC effect than participants primed with a difference-focus. These results can easily be interpreted in line with the *when* question. While 'keeping constant the pairings' (p.280), Corneille and colleagues solely manipulated the top-down input on the organism's side by changing its focus on the stimuli. A focus that fosters mental episodes by emphasizing the togetherness of stimuli (i.e., the similarity-focus) led to stronger EC effects than a focus that impeded mental episodes by emphasizing segregation of stimuli (i.e., the difference-focus).

Furthermore, Field and Moore (2005) showed that manipulating attention towards pairings influenced the magnitude of the EC effect. This finding is in line with the dual-force model we suggest. Attention is a force on the organism's side. If attention is weak, the bottom-up input must be stronger for the formation of mental episodes. If attention is strong, the environmental structure may be less critical. However, if there is no spatio-temporal proximity, it is unlikely that attention to stimuli alone leads to mental representations. If the stimuli are presented far apart, attention as a top-down process may not be sufficient. It is conceivable that other top-down processes must be involved (e.g., schemas, foreknowledge). Still, whether the lack of attention disrupts associative structures or somehow changes propositions is not part of the *when* perspective.

Memory. Mierop et al. (2020) investigated the influence of memory and cognitive load on EC. The authors showed that cognitive load impaired the EC effect even when participants correctly remembered the individual stimuli of a pair. This study may provide the most direct support for our *when* question suggestion because it showed that the correct retrieval of the CS-US relation (and not the individual stimuli of a pair) was necessary for EC to occur. Hampering the formation of a joint mental representation by introducing top-down input (i.e., cognitive load) reduced the EC effect. In summary, there are several studies already showing the influence of the organism on EC effects (see also Beckers et al., 2009; Blask et al., 2012; Brunstrom & Higgs, 2002; Dedonder et al., 2010; Gast & De Houwer, 2012; Green et al., 2021; Hughes, Ye, Van Dessel et al., 2019; Kattner, 2012; Kattner & Green, 2015; Laane et al., 2010; Moran et al., 2016; Olson & Fazio, 2001). However, this research primarily aimed to answer the *how* question but can be reinterpreted as evidence for the *when* question. We believe the present dual-force perspective presents a framework to explain *when* co-occurrences in the environment are encoded into the same or different mental episodes, creating a necessary condition for EC effects.

4.10.3 Additional Clarifications

Based on Bless and Fiedler's dual-force model (Bless & Fiedler, 2006; Fiedler, 2001), we suggested bottom-up and top-down influences that force CS and US into the same or separate mental episodes. However, we did not measure if and how top-down influenced the formation of mental episodes. Instead, we created experimental conditions to manipulate the organism's influence on CS and US, separating CS and US into different mental episodes (Experiment 1-3, 5-6) or forcing CS and US into the same mental episode (Experiment 4); thus, we aimed to manipulate the assumed psychological construct rather than to measure it. A task for the future will be to provide a measure of mental episodes independent of measuring EC effects themselves. However, an even more promising approach will be a testable taxonomy of manipulations that target mental episodes, an approach that we are currently following.

Further, the dual-force perspective implies that in the absence of strong top-down processes, repeated co-occurrences of CS and US in the environment force them into the same mental episode. In the present work, we tested two possibilities of how top-down influences interact with a constant bottom-up influence. However, we did not test two additional possibilities that remain open. First, there is the combination of weak regularities in the environment (i.e., low spatial or temporal proximity of stimuli) with weak top-down processes. We believe this is the least interesting combination, and it is a priori true for the extreme cases of the absence of both spatial and temporal contiguity. Second, we did not test the combination of weak bottom-up input (i.e., low spatial or temporal proximity of stimuli) and strong top-down input. Gast and De Houwer (2013) provide an example of such a case by showing EC effects based on instructions. Similar to Jones et al. (2009), the authors interpreted their data in terms of a *how* question. Nevertheless, we believe it is empirically possible to dissociate the *when* question from a propositional process explanation.

The potential dissociation has a strong theoretical basis, allowing to substantiate the dual-force model for EC. For example Wertheimer's (1938) Gestalt principles or Radvansky's (2012) event horizon model suggest manipulations that should force stimuli into the same or separate mental episodes, creating or preventing EC effects under otherwise identical conditions. Likewise, a conceptual fit between CS and US should facilitate their storage in the same mental episode (i.e., top-down input). We are currently conducting a series of such experiments to further emphasize the importance of top-down input for the formation of mental episodes. We are confident that the present perspective will lead to novel predictions and a better understanding of the EC phenomenon.

4.10.4 Implications for the EC Phenomenon

The present perspective has two implications. First, it implies a hydraulic view of both bottom-up and top-down forces on the *when* question's answer (i.e., same mental episode or not). If there is no active part on the organism's side beyond basic perceptual processes and the suggested Gestalt and memory principles, the environment must provide strong and clear regularities; the constituents of such a regularity for a pairing would, amongst others, be spatial closeness, temporal closeness, high frequency of repetitions, but also exclusiveness of the pairing (see Alves et al., 2020). However, the more attention is given to co-occurrences, the more psychologically distinct it is for the organism, and the more enriched it is by existing knowledge structures, the less need exists for bottom-up regularities. On the end of the spectrum, one might consider no input from the environment but a pure mental representation of the CS and the US, such as in 'mere thinking' effects and self-generated attitude changes (Koehler, 1991; Tesser, 1978). Most of the time, though, both bottom-up and top-down influences will interactively determine whether CS and US are encoded into the same mental episode.

Second, our approach also offers explanatory potential for the difficulty in finding subliminal EC effects (e.g., Högden et al., 2018). According to Dehaene et al., (2006), subliminal processing of stimuli is characterized by weak bottom-up input that aggravates the spreading of neural activation. Subliminal presentation of CS or US (i.e., >50 ms; Hofmann et al., 2010) constitutes such a weak bottom-up process. In line with our assumptions, Dehaene and colleagues (2006) indicate that subliminal processing may occur if top-down processes (e.g., attention) interact with the bottom-up input. Such top-down influences might be conscious (e.g., Theeuwes, 2018) or unconscious (e.g., Gaspelin & Luck, 2018). However, given the weak input on the stimulus site (e.g., Heycke, Aust, & Stahl, 2017; Stahl, Haaf, & Corneille, 2016) and the lack of attention on the observer's part (e.g., Högden et al., 2018), it seems to be significantly more challenging in a subliminal EC paradigm to encode CS and US into the same mental episode.

Similarly, incidental EC, as introduced by the 'surveillance paradigm' (Olson & Fazio, 2001), should lead to weak EC effects, as both bottom-up and top-down forces are weak. In this paradigm, participants must react to 'target' stimuli (i.e., Pokémon) presented on a screen. Between this presentation of target stimuli, 'distractor' stimuli (i.e., Pokémon that were not used as target stimuli; CSs) that do not require an action on the participant's part are displayed together with valenced words or images (USs). Thus, participants' attention is drawn to the target stimuli (i.e., away from the CS-US pair). Simultaneously, the "rapid, nonrhythmic stream of perceptual events" (Olson & Fazio, 2001, p. 414) could indicate a weak bottom-up process, thereby weakening or eliminating EC effects (Moran et al., 2021).

4.11 Conclusion

The dual-force perspective shifts the focus from the environmental regularities to the top-down input from the organism's side. The perspective thereby sheds new light on old debates (e.g., subliminal or incidental EC) but also suggests new research avenues (e.g., a taxonomy of parameters that influence mental episodes) and thereby provides a framework that allows new predictions and supplements the highly relevant question how people acquire their likes and dislikes with an answer to the question *when* people acquire their likes and dislikes.

Open practices

We provide the stimulus materials, the preregistrations, R codes, and raw data files on our OSF page (https://osf.io/57ux2/?view_only=b5f 058da184a4d24b2a951f44ef30ab9).

OSF link: https://osf.io/57ux2/?view_only=b5f058da184a4d24 b2a951f44ef30ab9

Funding

The present research was funded by the Deutsche Forschungsgemeinschaft DFG (DFG UN 273/5–2)

Chapter 5: Manipulating Top-down and Bottom-up Input

In the previous chapter, a dual-force perspective on EC was introduced. First evidence was provided for the influence of both, bottom-up and top-down input on EC. More precisely, it was shown that the encoding of CS and US into the same mental episode was interrupted by asking participants a question in-between stimulus presentation. This interruption was independent of question type (i.e., location vs. valence of the US). However, instructing participants that CS and US belong together forced the stimuli into the same mental episode despite an in-between question. Nevertheless, the finding that instructions can influence EC is not a new one. For example, Gast and De Houwer (2013; see also Gast & De Houwer, 2012) showed that informing participants about the appearance of a stimulus together with something positive or something negative, led participants to like the stimulus that they expected to be followed by something positive more compared to the stimulus they expected to be followed by something negative. However, this has been considered mainly as evidence for EC's underlying processes, but not for the question of *when* EC should occur. Yet, a more direct test of top-down influence on EC will be given in the following chapter. Additionally, the experiments presented in the previous chapter did not differ regarding their bottom-up input. That is, the bottom-up input between experiments was held constant. The next chapter thus additionally presents experiments that vary with regard to their bottom-up input.

This chapter is based on the following manuscript:

Sperlich, L. M., & Unkelbach, C. (2024). Gute Gestalt – Why top-down input moderates EC, but bottom-up input doesn't. University of Cologne, Social Cognition Center Cologne. (manuscript in preparation)

Please note that certain modifications were made to the headings, citation style, and formatting to align with the layout of this dissertation. No changes were made to the content of the article.

Gute Gestalt – Why Top-Down Input Moderates EC, but Bottom-up Input Doesn't

Abstract

Evaluative Conditioning (EC) is defined as the change in the valence of a conditioned stimulus (CS) due to its pairing with a positive or negative unconditioned stimulus (US; De Houwer, 2007). In EC experiments, a pairing is usually implemented by presenting CS and US in close spatio-temporal contiguity. However, previous research showed that bottom-up-driven implementations of pairings are not independent of top-down input (i.e. the organism). Based on these findings, we investigated top-down and bottom-up moderators of the EC effect. In four pre-registered experiments (N = 553), we aimed to show that the manipulation of top-down and bottom-up features leads to different strengths in the EC effect. In the first two experiments, we manipulated the top-down influence using pretested fitting (e.g., rabbit – egg) versus non-fitting (e.g., rowing boat – riot) CS-US pairings. We did not find the expected effect in this first Experiment. In the second Experiment, we could demonstrate that fitting stimuli elicit a stronger EC effect than non-fitting stimuli. In the following two experiments, we manipulated the bottom-up input using brackets (Experiment 3) and colors (Experiment 4). However, none of the experiments yielded significant results. This observation allows three interpretations: (a) the strong influence of contiguity as a bottom-up input hampers contribution from other manipulations, (b) the applied manipulations were not strong enough, or (c) the manipulations had unintended effects on CS and US.

Keywords: evaluative conditioning, attitude acquisition, evaluative learning, environment-organism interactions, dual-force model

5.1 Introduction

Human behavior is, to a large extent, guided by likes and dislikes (Walther & Langer, 2008). While some of our likes are genetically determined (e.g., Poulton & Menzies, 2002), most of our preferences are learned (e.g., Baeyens et al., 1990). One way to learn such preferences is via Evaluative Conditioning (EC). EC can be defined as the change in the liking of a conditioned stimulus (CS) due to its pairing with an unconditioned stimulus (US) of positive or negative valence. In other words, observing an attractive (CS) holding a cup of coffee of an unknown brand (CS) might make us like the brand more. Conversely, observing an unattractive person together with the brand should lead us to like the brand less.

The high relevance of EC has been shown in several studies. For example, EC was demonstrated in clinical settings (e.g., Franklin et al., 2017), and consumer research (e.g., Ingendahl, Vogel et al., 2023). Moreover, EC has been shown to activate racial prejudice (e.g., Olson & Fazio, 2006) and is argued to influence political attitudes (e.g., Walther et al., 2005). Finally, EC can explain real-world phenomena like the 'kill-the-messenger' effect (Manis et al., 1974), which states that the valence of a message (US) is transferred to the people who convey the message (CS). Similarly, it might explain persuasion, where the change in the valence of an initially neutral message (CS) in the direction of the evaluated communicator (US) can be observed (Walther & Lange, 2008). However, even though EC might have real-world implications, and EC effects are usually observed in laboratory settings (for reviews see e.g., Gast et al., 2012; Moran et al., 2023; for a meta-analysis see Hofmann et al., 2010), finding EC effects in the real world seems to be a rather rare event (e.g., De Houwer et al., 2001).

In the following, we will explain why EC might not be observed more frequently in the real world. To do so, we will refer to the dual-force perspective put forward by Sperlich and Unkelbach (2022): We will summarize the influences of bottom-up and top-down input on the construction of mental episodes and review the predictions made by the model. Based on these predictions, we will investigate the influence of existing knowledge structures (Experiments 1 & 2), and bottom-up input (other than spatio-temporal contiguity; Experiments 3 & 4) on the encoding of CS and US into the same mental episode. Finally, we will discuss why top-down input appears to influence the encoding of CS and US into a shared mental episode, whereas bottom-up input does not.

5.1.1 A Dual-Force Perspective on EC

In our complex learning environment, we are constantly exposed to neutral (CS) and valenced stimuli (US). Thus, one might assume that EC is constantly observed in our everyday life. However, if this was the case, we would eventually develop the same indifferent attitude towards all objects in our environment (Hütter et al., 2014). Hence, the question arises as to why EC seems to be a quite robust effect in the laboratory (e.g., De Houwer et al., 2001), but rarely occurs in real life. A possible answer to this question is put forward by Sperlich and Unkelbach (2022). The authors argue that EC experiments differ from the appearances of CS and US in our typical environment. This difference becomes apparent when one takes a closer look at the structure of a typical EC experiment.

In typical EC experiments, CS and US are presented in close spatio-temporal proximity (e.g., Kukken et al., 2019). This structure of a typical EC experiment ('CS-US, blank, CS-US, blank') mirrors the rhythmic stream known from Gestalt psychology's law of proximity ("tap-tap, pause, tap-tap, pause", Wertheimer, 1938, p. 74). As following such a rhythm is almost inevitable (Neisser, 2014), CS and US are perceived as belonging together and are thus encoded into the same mental episode. In previous research, this presentation of CS and US in close spatio-temporal proximity was accepted as the sufficient predeterminant of EC effects (e.g. Hütter et al., 2012; Jones et al., 2010). However, this assumption does (a) follow from a behavioristic tradition that neglected the organism's role for the construction of mental episodes, and (b) overlooked the question of *when* EC occurs. That is, regarding the organism's role, EC research focused on the question of *how* EC occurs.

The *how* question is concerned with the underlying processes of EC, that is "the question of which cognitive processes intervene between mere co-occurrence and attitude formation or change" (Jones et al., 2010, p. 205). The organism's role regarding the *how* question has been investigated in various studies (for reviews see e.g., Jones et al., 2010; Sweldens et al., 2014; Moran et al., 2023); for example, in the context of memory (e.g., Halbeisen et al., 2014), attention (e.g., Kattner, 2012) or awareness (e.g., Högden et al., 2018).

The *when* question, on the other hand, does not make assumptions about the underlying processes of EC. Instead, this question suggests that if CS and US are encoded into the same mental episode, any kind of process might explain *how* EC occurs. For example, CS and US might be linked via propositions (e.g., 'CS and US co-occur', 'objects that co-occur are similar', 'CS and US are of similar valence', De Houwer, 2018), or via simple associations ('CS – US', e.g., Baeyens et al., 1992).

However, the encoding of CS and US into the same mental episode is also dependent on the organism's input. As mentioned before, EC is defined as the change in the liking of a CS due to its pairing with a valenced US; that is due to the contiguity of CS and US. Because contiguity is usually defined in terms of the co-occurrence of stimuli in space and time (De Houwer et al., 2020), such a definition focuses on the influence of environmental input on EC. However, Sperlich and Unkelbach (2022) argued that such a definition might be too narrow. By manipulating the organism's influence, they demonstrated that spatio-temporal contiguity (i.e., bottom-up input) leads to the encoding of stimuli into the same mental episode; however, only if the organism's influence (i.e., top-down input) did not interfere. Thus, in addition to bottom-up input, top-down input has been shown to influence stimulus encoding and EC.

In terms of EC experiments, this explains why EC effects are constantly observed in the laboratory, but rarely so in our everyday life. In EC experiments, strong bottom-up input is combined with weak (interfering) top-down, thus resulting in EC effects. In the real world, however, the failure to observe EC is attributable to the combination of weak bottom-up *and* weak top-down input. Yet, strengthening the top-down input should support the encoding of CS and US into the same mental episode. Thus, top-down input should have an influence above and beyond spatio-temporal contiguity.

5.1.2 The Previous Research

The dual-force perspective introduced by Sperlich & Unkelbach (2022) assumes that such a top-down force might be an existing knowledge structure provided by the organism. The authors argue that pre-existing schemas can lead to EC; even if propositionally invalid. For example, advertisements, where diapers (CS) appear together with attractive parents (US) might lead to EC due to a conceptual fit between parents and diapers; and despite an invalid proposition (i.e., people who change diapers do usually not look happy). In these cases, a strong bottom-up input meets a strong top-down input, which should support the encoding of CS and US into the same mental episode.

The authors provided a first test of the assumed top-down influences by asking participants to make a location or valence judgment between the presentation of CS and US. They were able to demonstrate that a location or valence judgment between CS and US prevented the encoding of CS and US into the same mental episode, as reflected in a reduced EC effect. However, informing participants that the two pictures presented on a screen are a pairing re-established the mental episode, despite the in-between judgment.

A very similar finding was presented by Moran (2024). Participants who observed CS-US co-occurrences without any additional instruction showed standard EC effects. However, when implementing top-down input (i.e., information about the unrelatedness of CS and US), EC effects were reduced. Moran (2024) discusses her findings in terms of the *how* question of EC (i.e., the underlying processes). However, this result might also be interpreted in terms of the *when* question (i.e., the dual-force perspective). Informing participants about the unrelatedness of CS and US might disrupt their encoding into the same mental episode.

Yet, both studies primarily demonstrated the influence of instruction on EC; this finding is not a new one (e.g., Gast & De Houwer, 2013).

5.1.3 The Present Research

Thus, with the present study, we aim to provide a more direct test of the assumption put forward by Sperlich and Unkelbach (2022). More precisely, we want to examine whether fitting CS-US pairings (e.g., bride – suit) lead to stronger EC effects than non-fitting CS-US pairings (e.g., champagne – dolphins), based on existing knowledge structures (usually, the groom wears a suit at a wedding). In Experiment 1, the effect of fitting and non-fitting CS-US pairings is tested online. However, as online experiments might negatively affect the quality of data due to the participant's inattention (Newman et al., 2021), and because we used a stimulus set that was not pretested for its neutrality, Experiment 2 replicates Experiment 1 in the laboratory, adding a CS prerating.

Furthermore, the dual-force model predicts that additional bottom-up information should have an impact on stimulus encoding. That is, adding bottom-up input alongside spatio-temporal contiguity should further support the encoding of CS and US into a shared mental episode. Because this influence was not tested before, we introduce such bottom-up manipulations in the subsequent experiments. In Experiment 3, we use brackets (encompassing vs. separating the CS-US pair) to manipulate the bottom-up input; in Experiment 4, we implement colors as bottom-up input (same vs. different colors as stimulus backgrounds). We predicted that emphasizing the togetherness of CS and US influences EC effects above and beyond spatio-temporal contiguity.

In total, we conducted five experiments. However, we do not report the third experiment in this series for two reasons: (a) We preregistered a sequential analysis and aborted data collection after the analysis of the first set of data (with $\alpha = .0294$) because we considered the occurrence of an effect highly unlikely, and (b) the experimental structure differed significantly from the other experiments reported here. The experiment is presented in the supplement.

5.2 Transparency and Open Science

Predictions and sample sizes for all four experiments were pre-registered. Assuming a standard EC effect (Simonsohn, 2014), sample sizes were based on the average effect size reported by Hofmann and colleagues (2010). We conducted the first experiment online. We switched from online to laboratory after the first experiment to increase data quality. However, we had to switch back to online after Experiment 3 due to timing constraints. To save resources, we pre-registered Experiments 2-4 as sequential analysis, assuming that we might already be able to detect an effect after collecting data from 60 (Experiment 2), and 100 (Experiments 3 & 4) participants. We stopped collecting data after 60 participants in Experiment 2. We continued data collection for Experiments 3-4 and report the results for these increased sample sizes with adjusted α-values ($\alpha = .0294$). We computed a post-hoc p-curve analysis (see Simonsohn et al., 2015) for the interaction effects' F-values of the top-down manipulations of Experiments 3 and 4 as well as the Experiments reported by Sperlich and Unkelbach (2022). This analysis places our average power at 94 %.

We report all data exclusions (all pre-registered), all manipulations, and all measures. Pretest, preregistrations, data, analysis code and materials are available from the Open Science Framework (OSF; https://osf.io/2jzfk/?view_only=f85bbf1f539046bd93653c4f6756298c). The reported studies are in accordance with the German Society for Psychology's (DGPs) guidelines for research with human participants (§7.3), which are based on the APA's ethical principles (APA, 2002).

We used the R software (R Core Team, 2021) to analyze our data. The following R packages were used: afex (Singman et al., 2016), effsize (Torchiano, 2020), dplyr (Wickham et al., 2021), ggpubr (Kassambara, 2020), magrittr (Bache & Wickham, 2020), plyr (Wickham, 2011), psych (Revelle, 2021), reshape2 (Wickham, 2007), rio (Chan et al., 2023), stringr (Wickham, 2023), tidyr (Wickham et al., 2024), and vctrs (Wickham et al., 2023).

5.3 Experiment 1: Prior Knowledge (Online)

In Experiment 1, we used different picture stimuli to manipulate participants' top-down input. Based on a pretest, four CS-US pairings were assigned to depict non-fitting stimulus pairs (thus representing a typical EC experiment). Four different CS-US pairings were chosen to represent fitting CS-US pairs (e.g., suit – bride). The stimulus assignments did not change between participants. We expected a stronger EC effect for CSs paired with fitting USs (vs. CSs paired with non-fitting USs).

5.3.1 Method

Participants. As pre-registered, we collected data from 100 participants on Prolific[17], who participated for payment of £0.50 (about $0.63). No participant met the pre-registered exclusion criteria. Thus, data from all 100 participants were retained in the sample (61 women, 38 men, 1 non-binary; M_{age} = 39.50, SD_{age} = 14.26).

Materials. We compiled the stimulus material by selecting CS-US pairs that we deemed to fit (e.g., hearing aid – old sick man), or deemed to not fit (e.g., bike – old sick man). Thus, every CS was assigned to one 'fitting' US and one 'non-fitting' US. We pretested the material online with 20 participants. Participants were asked to indicate how well the presented pairs fit, using a Likert scale ranging from -50 ('not at all') to + 50 ('very much'). In the first step, we computed the mean for every 'fitting' and every 'non-fitting' CS-US pair. In the second step, we calculated the difference score between every 'fitting' and 'non-fitting' CS-US pair. In the third step, we chose the 16 CS-US pairs (four 'fit positive', four 'non-fit positive', four 'fit negative', four 'non-fit negative') with the highest difference score as CS-US pairs for our experiments. All fitting CS-US pairs were rated at a minimum of +10 on the Likert scale, while all non-fitting pairs were rated at a minimum of -10 (with exception of the 'hat' CS). The eight CSs were selected from internet sources (for details see OSF), the USs were taken from the International Affective Picture System[18] (IAPS, Lang et al., 2008). As positive USs, we used image 1610, 2209, 5910, 8080, 1920, 7502, 8370, and 8420. As negative USs, we used image 2910, 9000, 9340, and 9600. We programmed the experiment with Qualtrics (Qualtrics, Provo, UT).

Design. We varied fit and valence within participants, resulting in a 2 (conceptual fit: yes vs. no) x 2 (US valence: positive vs. negative) within-participants design.

---

[17] Prolific automatically continues collecting data when participants did not finish the study. Thus, out of 112 participants, data from 102 participants was collected (no data is available for 10 participants that returned the study). 2 of these remaining participants did not finish the study and were excluded by Prolific, leaving us with the preregistered 100 participants.

[18] We inadvertently used IAPS images even though they are not permitted with open-source survey use (like Prolific; see https://csea.phhp.ufl.edu/Media.html) for online research. However, Experiments 2 and 3 were conducted in the laboratory. For Experiment 4, we changed the images.

Procedure. The computer randomly selected two pairs of each valence (positive vs. negative) and each concept (fit vs. non-fit) for each participant, resulting in eight pairs per participant. As the assignment of each pair was predetermined, the CS 'eggs' was for example always presented with positive valence. Additionally, this CS was always presented with a rabbit in the *conceptual fit* trials and with a rafting boat in the *conceptual non-fit* trials[19]. The CS 'church window', on the other hand, was always presented with USs of negative valence (e.g., fit: graveyard; non-fit: violent gang). Each CS-US pairing appeared three times, resulting in 24 CS-US trials (12 fit trials, 12 non-fit trials). During the conditioning phase, CS and US were presented simultaneously in the center of the screen. The stimulus pair was presented for 1500 ms. After each pairing presentation, participants had to press a 'next' button to ensure that they did not leave their device during the presentation phase. After a blank of 500 ms, the next trial started. Following this conditioning phase, the measurement phase started. Participants were asked to rate the liking of the CSs on a scale ranging from -3 ('not at all') to +3 ('very much'). The rating order was randomized for each participant. After providing demographic information, they were debriefed and paid.

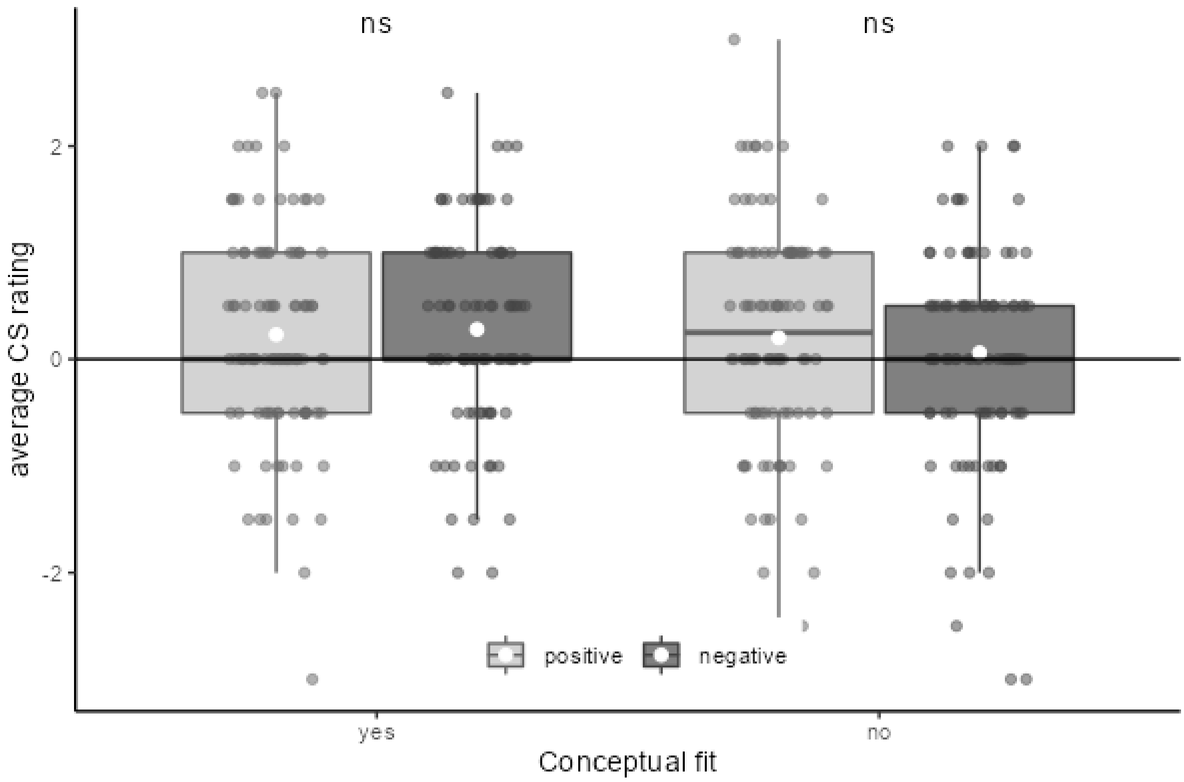
5.3.2 Results

We pre-registered a main effect for valence (i.e., a standard EC effect) and an interaction of valence and concept. That is, we expected the average ratings of CSs in the fit condition to be more extreme than in the non-fit condition. However, neither the main effect for valence, $F(1, 99) = 0.27$, $p = .607$, $\eta p^2 < 0.01$, nor the interaction were significant, $F(1, 99) = 1.31$, $p = .254$, $\eta p^2 = 0.01$. As Figure 5.3 shows, there was no difference between CSs presented with positive ($M = 0.22$, $SD = 1.03$) and CSs presented with negative USs ($M = 0.17$, $SD = 0.98$). Additionally, participants did not rate CSs presented with fit-positive USs ($M = 0.23$, $SD = 0.99$) more positive than CSs presented with nonfit-positive USs ($M = 0.20$, $SD = 1.08$). Similarly, average ratings of CSs presented with fit-negative USs ($M = 0.28$, $SD = 0.96$) were not significantly lower than CSs with nonfit-negative USs ($M = 0.06$, $SD = 0.99$).

[19] An overview of all stimulus pairings can be found on OSF.

Figure 5.3

Combined Box and Scatter Plots of Participants' CS Ratings in Experiment 1 as a Function of US Valence (Positive vs. Negative; Within Participants) and Conceptual Fit (Yes vs. No; Within Participants).



Note. Higher values indicate more positive CS evaluations. The box plots constitute the first quartile, third quartile, and the median. The circle within a box plot indicates the mean. Each dot represents a CS evaluation.

5.3.3 Discussion

In Experiment 1, we tried to manipulate participants' top-down input using fitting and non-fitting stimulus pairs. We expected pairs that conceptually fit to pronounce the EC effect. Such conceptual fits are for example used in the advertisement section: footballers advertise for sneaker brands, models praise healthy food, and race drivers promote fast cars. Yet, the implementation of such a conceptual fit in Experiment 1 did not prove successful. However, there are two straightforward explanations that might account for the failure to find the predicted results. First, in typical EC experiments, CSs like for example shapes (e.g., Gawronski et al., 2015), Chinese characters (e.g., Benedict et al., 2019), or unfamiliar brand names (e.g., Till & Priluck, 2000) are used. Usually, these stimuli are neutral in valence (Moran et al., 2023). However, in Experiment 1, the CSs were selected to enable a fit with certain USs. Thus, the CSs might not have been entirely neutral, leading to a distortion in the CS rating. Second, the experiment was conducted online. It has been argued that the quality of online data can be negatively influenced by inattentive participants (Newman et al., 2021). Hence, a second experiment was conducted to address these issues.

5.4 Experiment 2: Prior Knowledge (Laboratory)

Experiment 2 aimed to replicate Experiment 1, adopting two changes: (a) the data was collected in the laboratory and (b) participants were asked to prerate the CSs. Thus, we implemented a prerating before the presentation phase. Because the CSs were not taken from a database with pretested stimuli, the pretest was used to control for any pre-existing attitudes.

5.4.1 Method

Participants. We pre-registered a sequential analysis (cf. Lakens, 2014). As pre-registered, we collected data from 60 participants[20] on a University campus. Participants received 1 coffee coupon and 1 chocolate for their participation. As we found the pre-registered effect after collecting the first 60 participants using the adjusted α -value ($\alpha = .0294$), we stopped data collection at this point. No participant met any of the pre-registered exclusion criteria; thus, all participants were retained in the sample (30 women, 29 men, 1 other; $M_{age} = 23.13$, $SD_{age} = 6.75$)

Materials, Design and Procedure. We used the same material and design as in Experiment 1. The procedure was similar to the procedure in Experiment 1, with one exception: Before the acquisition phase started, participants were asked to rate the liking of the CSs on a slider[21] ranging from -3 (‘not at all’) to +3 (‘very much’). The aim of this prerating was to control for any pre-existing attitudes in the post-conditioning assessment. Thus, we calculated postrating-prerating for every CS

[20] 1 participant aborted the study before declaring consent and was thus replaced by another participant.
[21] We used a slider for the prerating and a Likert scale for the postrating of CSs to avoid response tendencies.

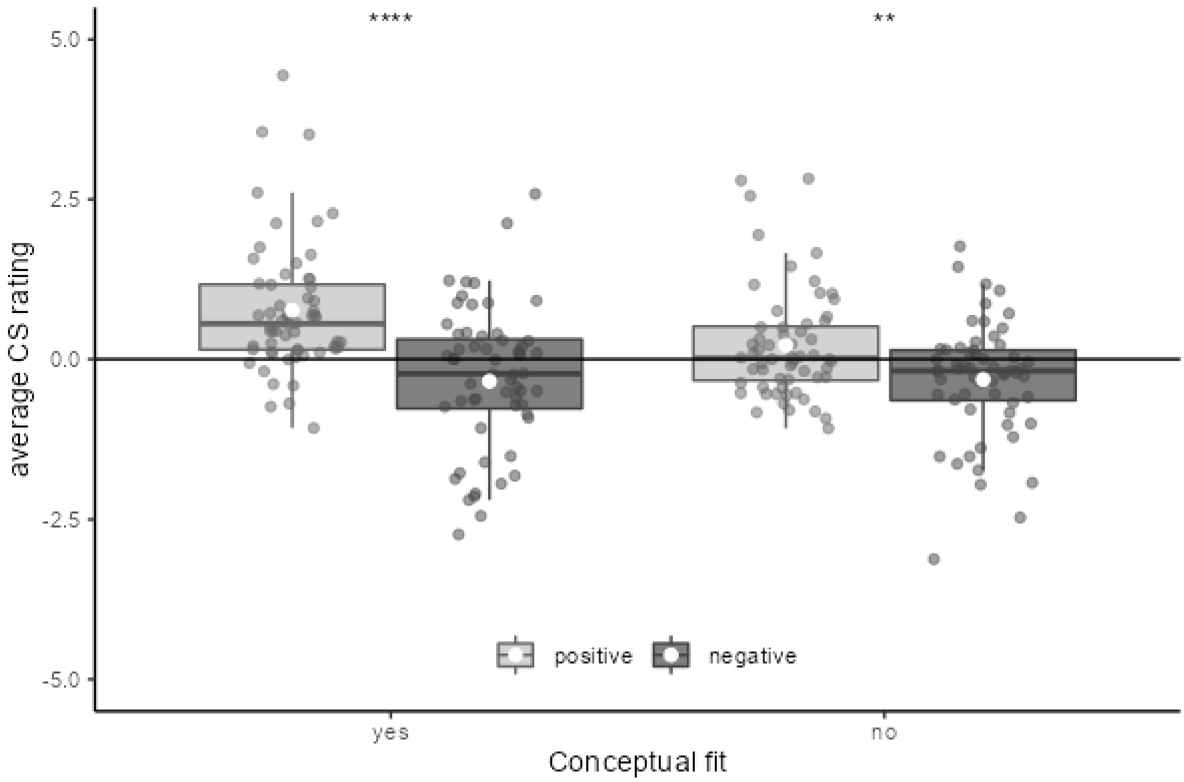
before analyzing the data. Additionally, we translated the instructions and labels from English to German.

5.4.2 Results

The ANOVA showed the predicted main effect for valence, $F(1, 59) = 21.75$, $p < .001$, $\eta p^2 = 0.27$. Participants rated positively paired CSs more positively ($M = 0.49$, $SD = 0.98$) than negatively paired CSs ($M = -0.33$, $SD = 1.06$). Additionally, we found the predicted interaction, $F(1, 59) = 8.08$, $p = .006$, $\eta p^2 = 0.12$. When paired with positive, fitting USs, CSs were rated more positively ($M = 0.76$, $SD = 1.02$) than when paired with positive, non-fitting USs ($M = 0.22$, $SD = 0.87$). Similarly, CSs paired with negative, fitting USs were rated more negatively ($M = -0.35$, $SD = 1.21$) than CSs paired with negative, non-fitting USs ($M = -0.32$, $SD = 0.90$). Figure 5.4 depicts this difference in average CS ratings.

Figure 5.4

Combined Box and Scatter Plots of Participants' CS Ratings in Experiment 2 as a Function of US Valence (Positive vs. Negative; Within Participants) and Conceptual Fit (Yes vs. No; Within Participants).



Note. Higher values indicate more positive CS evaluations. The box plots constitute the first quartile, third quartile, and the median. The circle within a box plot indicates the mean. Each dot represents a CS evaluation.

5.4.3 Discussion

Experiment 2 aimed to replicate Experiment 1, addressing two drawbacks of the first Experiment. First, Experiment 2 was conducted in the laboratory instead of online. Thus, poor data quality due to inattentive participants was minimized. Second, we asked participants to prerate the CSs before the presentation phase. By subtracting the postrating from the prerating, we could control for any pre-existing attitudes toward the CSs.

As predicted, there was a standard EC effect. More interestingly, however, we also found the predicted interaction. That is, participants rated CSs more positively (negatively) if paired with fitting vs. non-fitting USs. Even though we did not fully randomize the stimulus assignments, this pattern is difficult to explain without assuming a conceptual fit of some CS-US pairs.

However, so far, we only manipulated the top-down input. Yet, a manipulation of bottom-up input might also result in a pronounced EC effect (cf. Sperlich & Unkelbach, 2022). We tested this assumption in Experiment 3.

5.5 Experiment 3: Encompassing vs. Separating Brackets

In Experiment 3, we switched from manipulating the top-down input to manipulating the bottom-up input. Thus, instead of testing the conceptual fit between US and CS (top-down input), we implemented encompassing versus separating brackets between CS and US (bottom-up input). Because information within brackets usually belongs together, we assumed that brackets around CS-US pairings would emphasize the cohesiveness of CS and US. Conversely, brackets between CS and US should separate the stimuli. Additionally, we switched from a within-participants design to a mixed design. That is, the stimulus valence was manipulated within participants, the bracket assignment between participants.

5.5.1 Method

Participants. As in Experiment 2, we pre-registered a sequential analysis (cf. Lakens, 2014). We slightly increased the sample size due to the change in design. Thus, we collected data from the first 100 participants. As we did not find the effect for the predicted effect with $\alpha = .0294$, we continued data collection. We collected data from 204 participants[22] on a University campus, who received one chocolate and one coffee coupon for their participation. As pre-registered, five participants were excluded who requested not to use their data. Two participants were excluded because they were too young. Thus, 197 participants were retained in the analysis (119 women, 75

[22] We pre-registered 200 participants, but the research assistant's collected data from 204 participants.

men, 2 non-binary, 1 prefer not to say; M_{age} = 25.99, SD_{age} = 10.25). 102 participants were assigned to the 'separating brackets' condition and 95 participants to the 'encompassing brackets' condition.

Materials. We used the same materials as in Experiments 1 and 2. However, different from previous experiments, we did not use the fitting CS-US pairs for Experiment 3. To avoid confounds with top-down input, only the non-fitting pairings were used (e.g., church window – violent gang). Thus, the positive USs set contained image 1920, 7502, 8370, and 8420; the negative set consisted of image 2141, 2683, 6821, and 9830. The experiment was programmed with OpenSesame (Mathôt et al., 2012).

Design. Valence (positive vs. negative) and brackets (brackets vs. no brackets) were manipulated within participants; type of bracket (encompassing vs. separating) was manipulated between participants. To analyze the data as preregistered, we first computed the post-prerating for every CS (cf. Experiment 2). We then averaged the ratings for every participant on the two within-levels: valence (positive vs. negative) and presence of brackets (yes vs. no). We computed the EC effect for the brackets within condition by subtracting the ratings for negatively from the ratings for positively paired CSs. Subsequently, we calculated the difference between the EC effect of brackets and no brackets within each condition. For the between-condition 'separating brackets', we predicted a larger effect for CS-US pairings presented without brackets than for pairings presented with brackets. Thus, the calculated difference should be negative. For the between-condition 'encompassing brackets', we hypothesized a stronger EC effect for pairings presented with brackets compared to pairings presented without brackets. Hence, the calculated difference should be positive. To test whether this difference is significant, we used a Welch Two Sample t-test[23].

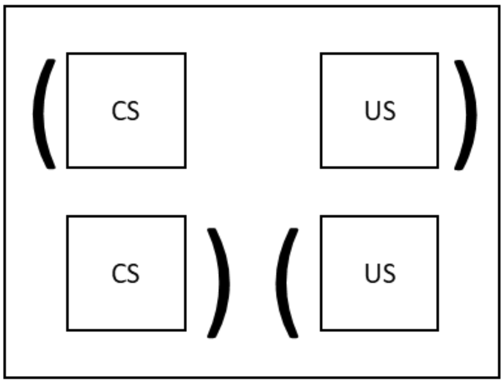
Procedure. The procedure was similar to the procedure of Experiment 2. Like in Experiment 2, participants observed two stimuli (CS and US), that were presented simultaneously. To avoid

[23] The usage of Welch's test instead of Student's t-test is recommended (see http://daniellakens.blogspot.com/2015/01/always-use-welchs-t-test-instead-of.html). Any fractional degrees of freedom are due to the Welch's test pooled degrees of freedom.

confounds with top-down input, we only used the non-fitting CS-US pairings from Experiment 2. The presentation order of the eight CS-US pairings was randomized. The computer program counterbalanced the bracket assignment for each participant. Two of the positive CS-US pairings and two of the negative CS-US pairings were presented with a bracket encompassing the pair ('encompassing brackets' condition) or a bracket between the pair ('separating brackets' condition; see Figure 5.5A).

Figure 5.5A

Schematic Illustration of the Encompassing and Separating Brackets.



Note. Spatio-temporal contiguity was held constant between unconditioned stimulus (US) and conditioned stimulus (CS). In the 'encompassing brackets' condition, the brackets were presented around CS and US (top), in the 'separating brackets' condition, the brackets were presented in-between CS and US (bottom).

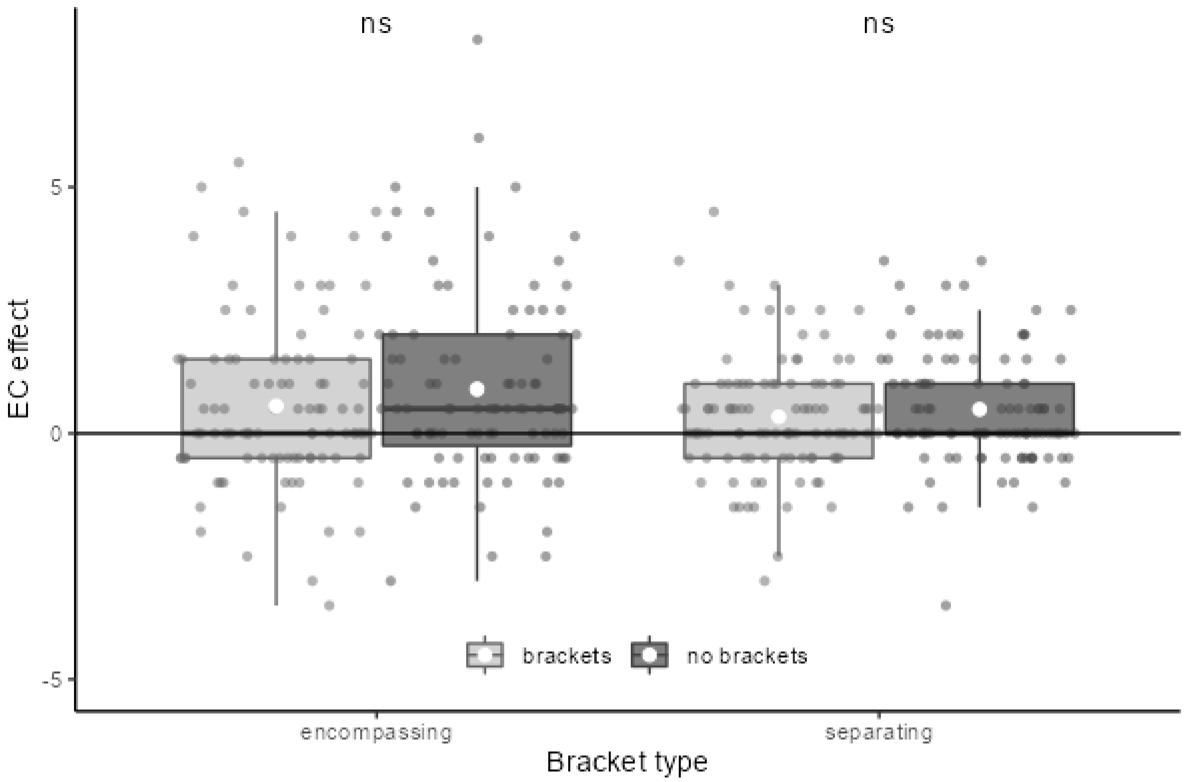
In both conditions, the other four CS-US pairings (two positive, two negative) were presented without brackets. Each CS-US pairing was presented for 2500 ms, followed by a blank screen (1500 ms). Different from Experiment 2, each pairing was presented six times, resulting in 48 CS-US presentations. Following the postrating, participants were asked to rate how much two pictures presented within brackets ('encompassing brackets' condition) or outside of brackets ('separating brackets' condition) belong together on a scale from 1 ('not at all') to 7 ('very much'). This question served as a manipulation check. After providing demographic data, participants were debriefed and received compensation.

5.5.2 Results

We predicted EC effects to be stronger for the 'encompassing brackets' than for the 'separating brackets' condition. To test this assumption, we used a Welch Two Sample t-test. Contrary to our prediction, the difference between conditions was not significant, $t(167.58) = 0.74$, $p = .458$, $d = 0.11$. The difference between the bracket-within condition (brackets vs. no brackets) was negative for both, the 'encompassing brackets' ($M = -0.37$, $SD = 1.41$) and the 'separating brackets' condition ($M = -0.15$, $SD = 2.00$). The EC effect across conditions was highly significant, $t(196) = 6.31$, $p < .001$, $d = 0.45$. Figure 5.5B shows the EC effect (i.e., ratings for positively - ratings for negatively paired CSs) for each condition. In the 'encompassing brackets' conditioning, EC was higher for CS-US pairings presented without brackets ($M = 0.89$, $SD = 1.89$) than for pairings presented with encompassing brackets ($M = 0.56$, $SD = 1.70$). Yet, this difference was not significant, $t(94) = 1.63$, $p = .105$, $d = 0.19$. Similarly, in the 'separating brackets' condition, EC was higher for pairings without brackets ($M = 0.49$, $SD = 1.17$) than for pairings with brackets ($M = 0.34$, $SD = 1.21$); however, this difference was also not significant, $t(101) = 1.08$, $p = .279$, $d = 0.13$.

Figure 5.5B

Combined Box and Scatter Plots of Participants' CS Ratings in Experiment 3 as a Function of US Presence of Brackets (Brackets vs. No Brackets; Within Participants) and Type of Bracket (Encompassing vs. Separating; Between Participants).



Note. Higher values indicate higher EC effects. EC effects were computed by subtracting the ratings for negatively paired CSs from the ratings for positively paired CSs of the 'brackets' and 'no brackets' trials, averaged for each participant. Each dot represents an averaged evaluation. The box plots constitute the first quartile, third quartile, and the median. The circle within a box plot indicates the mean.

As pre-registered, we also used a Welch Two Sample t-test to test for differences between CS-US pairings presented with brackets and CS-US pairings presented without brackets in both conditions. Neither the test for the ‘brackets’ comparison was significant, $t(168.84) = 1.03$, $p = .303$, $d = 0.15$, nor was the test for the ‘no brackets’ comparison, $t(154.54) = 1.79$, $p = .075$, $d = 0.26$.

Additionally, we analyzed the manipulation check data. Ratings in the ‘encompassing brackets’ condition should be higher (i.e., pictures within brackets belong together) than ratings in the ‘separating brackets’ condition (i.e., pictures outside of brackets do not belong together). However, there was no difference between the ratings, $t(194.15) = -0.65$, $p = .518$, $d = -0.09$. Participants rated the cohesiveness of CS and US within a bracket equally low ($M = 3.00$, $SD = 1.80$) as outside of brackets ($M = 3.16$, $SD = 1.81$).

5.5.3 Discussion

With Experiment 3, we wanted to test the influence of bottom-up input on the encoding of CS and US as a pairing. We hypothesized that a bracket around CS and US would support the encoding of the stimuli into the same mental episode, whereas separating brackets would prevent such encoding. Even though there was a basic EC effect, we did not find the predicted pattern. In each condition, the EC effect was descriptively larger for the CS-US presentations presented without brackets. None of the pre-registered effects reached significance.

However, the manipulation check showed that participants on average did not perceive CS and US within brackets differently from CS and US outside of brackets. Thus, the manipulation might not have worked as intended. As Figure 5.5A shows, the ‘separating brackets’ condition represents a pattern that is not typically encountered in everyday life. Our manipulation might have been stronger with brackets around the individual stimuli: (CS) (US). We did not use such a manipulation because the conditions were intended to be symmetrically structured. However, such an ‘asymmetrical’ structure might have elucidated that CS and US depict two different pieces of information, thereby emphasizing the cohesiveness of CS and US presented within the same bracket.

Still, in scientific papers and other literature, brackets are usually used to communicate additional information that is not inherently relevant for understanding. Conversely, CS-US pairings presented without brackets might be perceived as the important part (compared to the information within brackets), resulting in a (descriptively) stronger EC effect. Thus, in Experiment 4, we changed the bottom-up information to color.

5.6 Experiment 4: Same versus Different Colors

In contrast to brackets, colors have been used in EC experiments before. Using a variant of the Eriksen flanker task, Blask and colleagues (2017a) for example could show that congruent trials (i.e., CS and US were presented with frames of the same color) elicited stronger EC effects than incongruent trials (i.e., CS and US were framed with different colors). Similarly, Hughes and colleagues (2020) demonstrated that when a target object (e.g., a CS) is simultaneously presented together with both, a positive and a negative source object (e.g., a US), the target object acquires the valence of the source object with which it shares a feature (e.g., color). Additionally, the color manipulation is similar to the 'common region' proposition by Palmer (1992): "Elements will be perceived as grouped together if they are located within a common region of space, i.e., if they lie within a connected, homogeneously colored or textured region or within an enclosing contour" (p. 438). Thus, in Experiment 4, we used same versus different colors as bottom-up input. Due to timing constraints, we conducted Experiment 4 online.

5.6.1 Method

Participants. We collected data from 200 participants via Prolific[24], who received £1.05 (approx. $1.33) for their participation. As preregistered, we run a sequential analysis on 100 participants with α = .0294. Because we did not find the predicted effect, we continued data collection (cf. Lakens, 2014) until we obtained the full sample. As preregistered, four participants were excluded who reported to not have perceived the colored backgrounds. Thus, 196 participants were retained in the analysis (114 women, 81 men; 1 non-binary, M_{age} = 42.82, SD_{age} = 13.35).

[24] As in Experiment 1, Prolific automatically continued collecting data when participants did not finish the study. Thus, out of 210 participants, data from 200 participants was collected (no data is available for 10 participants that returned the study).

Material, design and procedure. Due to timing issues, we switched back to an online setting. Because the usage of IAPS images is not permitted for online research with open-source surveys (like Prolific; see https://csea.phhp.ufl.edu/Media.html), we compared the USs used in Experiment 3 with pictures of the Open Affective Standardized Image Set (OASIS; Kurdi et al., 2017) and selected pictures that were comparable with the IAPS pictures in appearance and valence. The set of positive stimuli retrieved from OASIS contained Penguins 2, Rafting 2, Rollercoaster 1, and Happy face 1; the set of negative stimuli consisted of Garbage dump 2, Boxing 2, Police 2, and Depressed pose 1.

The experiment was programmed using Qualtrics (Qualtrics, Provo, UT). Both valence and color were manipulated within participants, resulting in a 2 (valence: positive vs. negative) x 2 (color: same vs. different) design. We predicted a main effect for valence (i.e., a standard EC effect). Additionally, we predicted an interaction for valence and color (i.e., a higher EC effect for the 'same color' condition than for the 'different color' condition) and preregistered to test this assumption with an ANOVA. The procedure was similar to the procedure of Experiment 3. However, we dropped the manipulation check and increased stimulus presentation to 3000 ms. Yet, the most important change was the usage of colors as bottom-up input[25]. The computer program counterbalanced the color assignment for each participant. Four CS-US pairings (two positive, two negative) were presented with the same background color for CS and US, the other four pairings (two positive, two negative) were presented with different background colors for CS and US. The assignment of different background colors was randomized. The boxes for background colors were larger than the boxes for the pictures. This ensured that background colors were clearly visible. Everything else was identical to Experiment 3.

5.6.2 Results

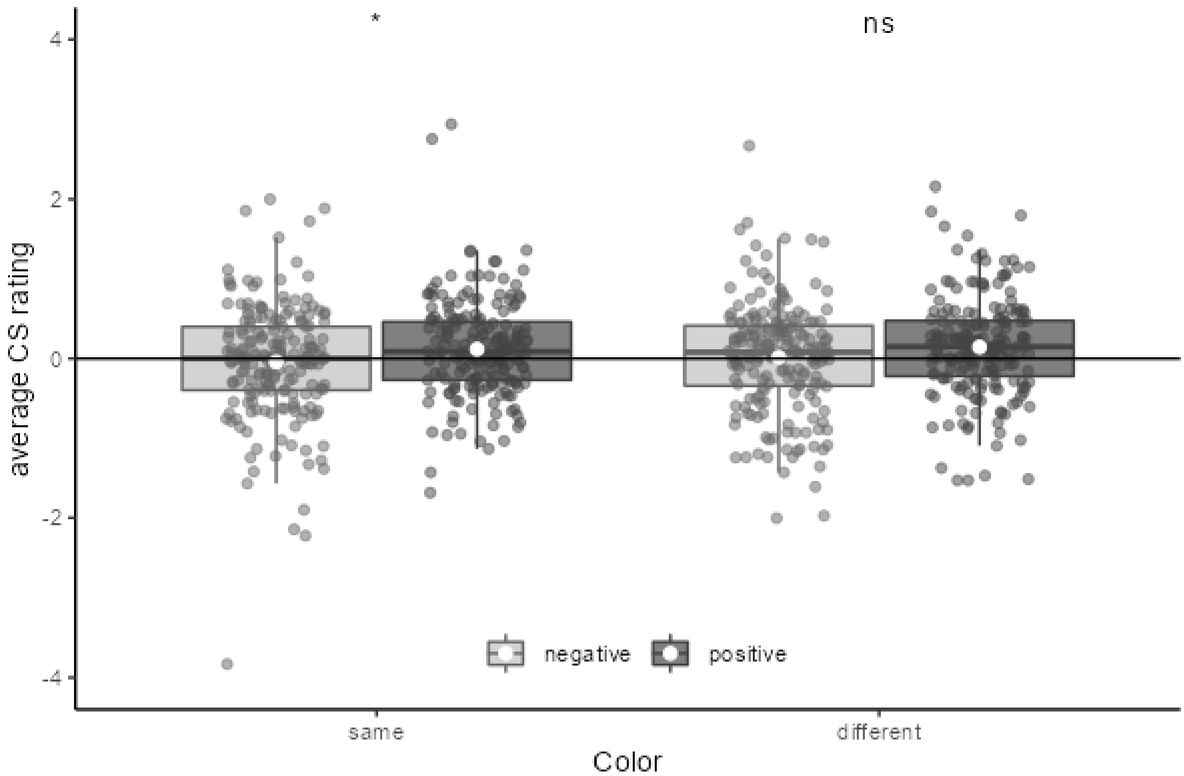
We predicted a larger EC effect for the 'same color' than for the 'different color' condition. We expected CSs presented with USs of the same background color to be rated more extremely than CSs whose background color differed from the background color of USs. Whereas we found the main

[25] An overview of all colors used in Experiment 4 is uploaded on OSF.

effect for valence (i.e., a standard EC effect), $F(1, 195) = 7.34$, $p = .007$, $\eta p^2 = 0.04$, we did not find the predicted interaction effect, $F(1, 195) = 0.11$, $p = .745$, $\eta p^2 < 0.01$ (see Figure 5.6). Yet, in the 'same color' condition, CSs presented with positive USs were rated on average more positive ($M = 0.12$, $SD = 0.61$) than CSs presented with negative USs ($M = -0.04$, $SD = 0.72$).

Figure 5.6

Combined Box and Scatter Plots of Participants' CS Ratings in Experiment 4 as a Function of US Valence (Positive vs. Negative; Within Participants) and Color (Sames vs. Different; Between Participants).



Note. Higher values indicate more positive CS evaluations. The box plots constitute the first quartile, third quartile, and the median. The circle within a box plot indicates the mean. Each dot represents a CS evaluation.

Similarly, in the ‘different color’ condition, participants rated CSs presented with positive USs on average slightly more positive (*M* = 0.15, *SD* = 0.62) than CSs presented with negative USs (*M* = 0.01, *SD* = 0.69). The main effect for color was not significant, $F(1, 195) = 1.26$, $p = .262$, $\eta p^2 < 0.01$.

To further explore the data, we used a one-way ANOVA to test the difference between positively and negatively paired CSs within each condition. Using a one-way ANOVA, differences between valence proved to be significant in both, the ‘same color’, $F(1, 195) = 4.83$, $p = .029$, $\eta p^2 = 0.24$, and the ‘different color’ condition, $F(1, 195) = 4.34$, $p = .038$, $\eta p^2 = 0.22$. Yet, this analysis was done post-hoc; that is, we did not preregister the analysis.

5.6.3 Discussion

Based on the findings of previous EC experiments and the suggestion by Palmer (1992), we changed the bottom-up input in Experiment 4 to same versus different background colors of CS and US. We assumed that presenting CS and US with the same background color would support the encoding of CS and US into the same episode, whereas presenting CS and US with different background colors would hamper such encoding. Thus, we predicted larger EC effects for the former compared to the latter condition. Using an explorative analysis, a small difference between conditions was observable. That is, the EC effect was more pronounced in the ‘same color’ compared to the ‘different color’ condition. Yet, we did not find this difference using the preregistered test. Because the assignment of colors was randomized, it is possible that some of the participants received rather similar colors in the ‘different color’ condition. We for example used #1560bd (a darker blue) and #76d7ea (a lighter blue) as different background colors. Yet, these ‘different’ backgrounds might be perceived as rather similar (Özgen, 2004), at least in the context of other background colors (e.g., pink and orange). Additionally, the random assignment of colors might have led to background presentations of complementary colors in the ‘different color’ condition. While the complementary feature should enhance the contrast between the colors, people might have prior knowledge about complementary colors appearing together (e.g., in clothing or paintings). Thus, in some cases, the manipulation might have counteracted the hampering feature of the ‘different color’ condition.

Another possibility is that the color manipulation was generally too subtle. In terms of Gestalt psychology, spatio-temporal proximity might already lead to the perception of CS and US as a unity, and thus, adding (same) colors may not have contributed further to the encoding of the stimuli into the same mental episode. This would explain why we find a standard EC effect, but no moderation. Yet, despite these post-hoc explanations, the findings are not compatible with the dual-force perspective introduced by Sperlich and Unkelbach (2022) without making additional assumptions.

5.7 General Discussion

While EC has been shown to be a rather robust effect in various laboratory settings (e.g., Fulcher & Hammerl, 2001; Kattner, 2012; Tello et al., 2018; Moran, 2024; for a meta-analysis see Hofmann et al., 2010), other researchers doubt its robustness (e.g., Jones et al., 2010; De Houwer et al., 2005; Martin & Levey, 1978); among other reasons, due to the lack of evidence of EC in real-world settings (Rozin et al., 1998). Yet, it was argued that experimental and real-world settings differ significantly (Sperlich & Unkelbach, 2022). In EC experiments, the encoding of CS and US into the same mental episode might be a by-product of the experimental settings that follow a determined rhythm (i.e., 'tap-tap, pause'). Because following such a rhythm is inevitable (Neisser, 2014), CS and US are perceived as a pairing in typical EC experiments. In the real world, however, this determined rhythm might be observed less frequently. To compensate for weak bottom-up input, stronger top-down input is required to encode CS and US into a shared mental episode. Conversely, if such top-down input interferes with strong bottom-up input, encoding should be hampered. First evidence for the importance of top-down input for stimulus encoding was provided by Sperlich and Unkelbach (2022).

We tried to extend these findings by demonstrating that existing knowledge structures (e.g., conceptual fit) have a supportive influence on bottom-up input (Experiments 1 & 2). Additionally, we aimed to provide evidence for further bottom-up input to influence stimulus encoding (Experiments 3 & 4).

We varied the top-down input by using CS-US pairings with conceptual fit (e.g., church window – graveyard) or non-fit (e.g., church window – violent gang). Although we did not find an effect of top-down input in Experiment 1 (online setting), the effect was significant in Experiment 2 (laboratory setting). In Experiment 3 and 4, we manipulated bottom-up input using brackets (Experiment 3) and colors (Experiment 4). Even though we found a significant EC effect across conditions in both experiments, the preregistered analyses did not yield significant results. However, the explorative investigation of EC effects in Experiment 4 showed a slightly more pronounced EC effect for CS-US pairings of the 'same color' compared to the 'different color' condition, suggesting that, if anything, further bottom-up input adds little to the high impact of spatio-temporal contiguity.

Taken together, our findings provide mixed evidence regarding the dual-force perspective (cf. Sperlich & Unkelbach, 2022).

5.7.1 (The Lack of) Evidence for the Dual-Force Perspective

Based on the dual-force perspective, supporting bottom-up as well as supporting top-down input should lead to the encoding of CS and US into the same mental episode. In terms of the dual-force perspective, the conceptual fit between CS and US represents existing knowledge structures, that should support the encoding of CS and US into the same mental episode. Similarly, encompassing brackets and a shared background color for CS and US should support such encoding.

The Influence of Top-Down Input. Although Sperlich and Unkelbach (2022) could show that judgments in-between CS-US presentations prevent the encoding of stimuli into the same episode, they only provided partial evidence for the supporting influence of top-down input. That is, if participants were instructed to observe "pairs of pictures" (vs. "single pictures"; p. 11), they encoded CS and US into the same mental episode regardless of the judgment in-between CS and US presentation. Still, it might be argued that this instruction triggers propositional processes and thus, might be better explained in terms of the *how* instead of the *when* question (but see Sperlich & Unkelbach, 2022, for a counterargument). Nevertheless, such an argument is more difficult to apply to the manipulation of the top-down inputs in Experiments 1 and 2. Conversely, the larger EC effect in

Experiment 2 is readily compatible with the dual-force perspective. By enriching co-occurrences with existing knowledge structures (e.g., ‘Easter bunny, Easter eggs’), the likelihood of CS and US being encoded in the same mental episode increases. The idea that learning improves when stimulus information is actively generated and enriched by prior knowledge was already proposed by Fiedler and Unkelbach (2011). The authors suggested that the CS-US relation must be “supported by appropriate encoding operations” for “conditioning to be effective” (p. 640). Thus, it was argued that encoding CS and US as *friends* would lead to EC effects while encoding CS and US as *enemies* would reverse this effect. Results showed that CS ratings were more positive (encoding scheme: friend) or more negative (encoding scheme: enemy) for participants who indicated high likelihood for CS and US to be friends or enemies, respectively. Thus, a pronounced (reversed) EC effect was found for participants who successfully generated assimilative (contrastive) information. In line with these findings, participants in Experiment 2 who actively retrieved existing knowledge structures to generate the link between, for example, rabbit and eggs, showed larger EC effects than participants who did not. However, while Fiedler and Unkelbach (2011) interpreted their findings in terms of the *how* question, the dual-force perspective does not make assumptions about whether the link between US and CS is of propositional or associative nature. A proposition such as ‘The Easter Bunny brings Easter eggs’ might be formed, or a simple associative link such as ‘Easter Bunny - Easter Eggs’.

The Influence of Bottom-Up Input. While there is some evidence suggesting that top-down input influences the encoding of CS and US into the same mental episode, the influence of bottom-up input is less clear. The bottom-up input in EC experiments is usually kept constant. This sequence of CS and US in EC experiments reflects the auditory organization previously proposed by Wertheimer (1938): tap-tap, pause. Thus, EC experiments adhere to a Gestalt psychological principle of proximity, stating that stimuli presented in this manner are perceived as belonging together. The importance of spatio-temporal proximity for EC experiments has been demonstrated before (e.g., Jones et al., 2009; Hughes, Mattavelli et al., 2018; Rydell & Jones, 2009). However, the dual-force perspective assumes that in addition to such spatio-temporal proximity, other bottom-up factors might further strengthen the encoding of CS and US into the same mental episode. Encompassing brackets or the same-colored background of CS and US should depict such factors. However, based on the preregistered analyses for Experiments 3 and 4, this assumption cannot be maintained. Neither brackets nor colors seemed to have an impact on encoding. This result is even more surprising when considering findings from other research. Both, Blask et al. (2017a) and Hughes et al. (2020) for example found an effect of color on stimulus ratings. Blask and colleagues (2017a) could show that presenting CS and US in same-colored frames pronounced EC compared to their presentation in different-colored frames. Similarly, Hughes and colleagues (2020) presented a CS together with two USs; one US shared a background color with the CS, the other did not. The authors found a valence transfer of the CS in direction of the US it shared a color with. Yet, the study from Hughes et al. (2020) crucially differed from Experiment 4 in at least one point: instead of two stimuli (as in typical EC experiments), three stimuli appeared on the screen. The CS was presented on the left side of the screen, while the USs were presented on the right side. Even though the authors attributed the valence transfer to the shared feature principle, it is conceivable that the encoding of the same-colored stimuli occurred based on the contrast with the other, differently-colored stimulus (cf. Bless & Schwarz, 2010). Additionally, the results of Blask et al. (2017a) are less straightforward than depicted. In another, very similar study, Blask et al. (2017b) again showed an EC effect for CSs and USs presented in same color frames; however, they also found an EC effect when CS and US were framed by different colors, which they explained with the attention-grabbing power of affective stimuli. However, because they presented CS and US in close spatio-temporal proximity, another explanation is conceivable. Perhaps spatial-temporal contiguity is such a strong influencing factor that further manipulations contribute little additional value as bottom-up input.

Top-Down vs. Bottom-Up. An organism does not have unlimited but at least has many degrees of freedom to construe the social world (e.g., Kunda, 1990). In line with this assumption, moderating the top-down influence (i.e., the organism’s influence) on co-occurrences seems to be possible. Sperlich and Unkelbach (2022) provided first evidence that presenting judgments in-between stimulus presentation prevents the encoding of CS and US into the same episode, whereas a pairing instruction re-establishes the mental episode. The present research adds to this finding by showing that conceptual fit influences stimulus encoding. Even though this finding highlights the importance of the organism's role for EC, the impact of bottom-up input is less clear. None of the experiments manipulating the bottom-up input showed a significant result. However, even though not directly compatible with the dual-force perspective, it is conceivable that additional bottom-up input, presented alongside spatio-temporal contiguity, does not have any further influence on stimulus encoding. Put differently, without interfering top-down input, everything that is presented "close together goes together" (Schneider & Mattes, 2021, p. 231). The authors noted that proximity is such a prominent factor in human perception that it takes effort to not perceive the following stream of dots ●● ●● as two dot pairs instead of four dots.

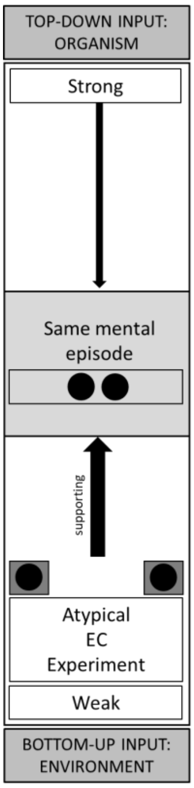
This assumption also fits with the EC effect across conditions found in both experiments manipulating bottom-up input (i.e., Experiments 3 & 4). Additionally, this assumption fits with the result by Moran (2024) mentioned before. Whereas Moran (2024) considers her findings not to be perfectly compatible with any of EC's process theories, they are readily explainable by adapting a dual-force perspective. By informing participants about the unrelatedness of CS and US, encoding might be disrupted. However, because CS and US are presented in close spatial-temporal proximity, bottom-up and top-down inputs work against each other. As a result, CS and US might sometimes not be encoded into the same mental episode, which reduces the EC effect but does not make it disappear completely.

Yet, while strong bottom-up input might interfere with top-down input, this effect seems to be pronounced for further bottom-up input. Experiments 3 and 4 may therefore not provide a fair test for the influence of other bottom-up input on the encoding of stimuli. A fairer test to investigate the influence of bottom-up input other than spatio-temporal contiguity might be to present such information (e.g., colors) in the absence of spatio-temporal contiguity. Thus, when discarding the strong spatio-temporal manipulation, the influence of other bottom-up manipulations might become apparent (cf. Figure 5.7A).

Still, one might argue that a color manipulation does not really depict bottom-up influence, but rather top-down input. Similar to the conceptual fit, the grouping of stimuli based on color might rely on past experience (Wertheimer, 1938) and thus, on existing knowledge structures. However, as colors can be described in terms of physical properties, they serve as environmental cues just like spatio-temporal contiguity. Additionally, the perception of colors does not reflect an internally generated response (cf. Fiedler & Unkelbach, 2011), but might represent a rather passive perceptual process and thus, constitute a weak top-down influence (cf. Sperlich & Unkelbach, 2022).

Figure 5.7A

Schematic Illustration of Bottom-Up Influence in the Absence of Spatio-Temporal Proximity.



Note. The size of the arrows represents the respective force's strength. The black dots represent CS and US. The dark grey background represents a shared background color of CS and US. Adapted from Sperlich & Unkelbach (2022).

5.7.2 Limitations and Open Questions

Even though we did not test the post-hoc explanation given in the previous section, it fits with our data. Yet, a test of the post-hoc explanation would not only provide the required evidence for this post-hoc assumption but also provide an environment that mirrors the real world. In the real world, two stimuli presumably rarely occur the way they do in EC Experiments. Thus, in everyday life, people do not encounter the same rhythm of stimuli that is provided in an experimental setup (i.e., 'tap-tap, pause'). Consequently, increasing the distance between US and CS might represent a scenario with higher external validity.

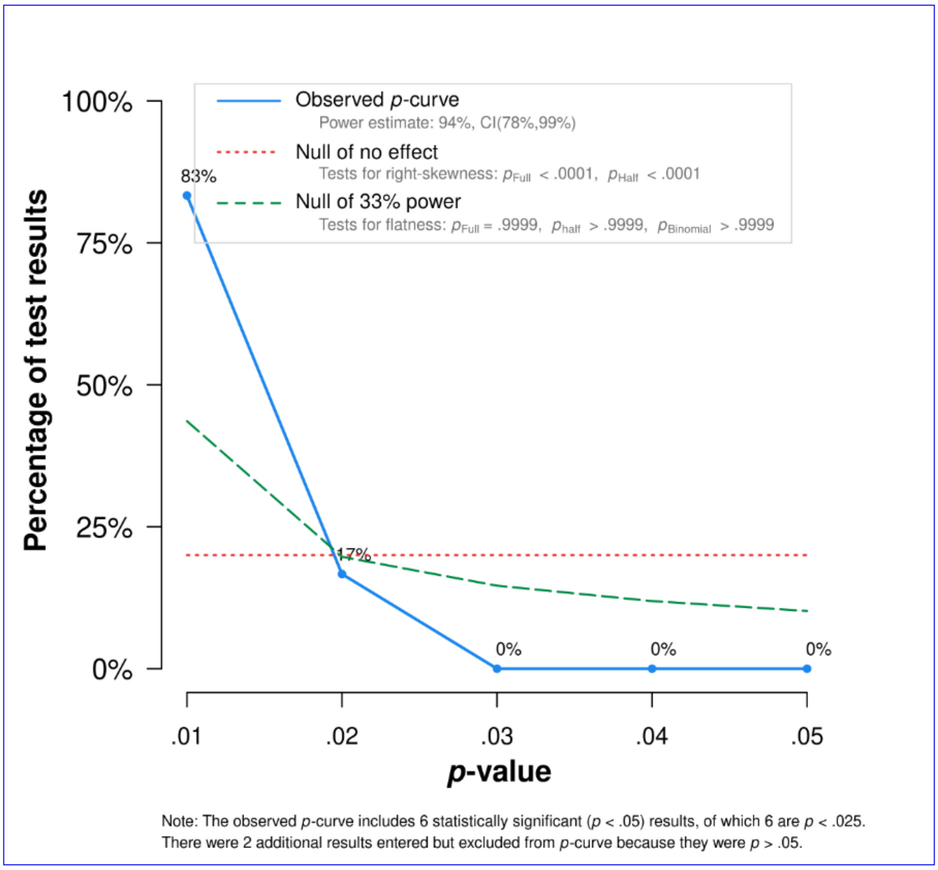
However, the present findings are not readily compatible with the dual-force perspective, which indicates an influence of bottom-up input in the presence of spatio-temporal contiguity. Additionally, four out of five experiments (including the experiment reported in the supplement) did not yield a significant result. This data might thus speak against the model. However, the insignificant result of Experiment 1 might be attributed to the choice of initially valenced CSs. After eliminating this confound, we found the predicted pattern. Still, we underestimated the influence of spatio-temporal proximity as a factor that influences the encoding of CS and US into a shared mental episode. In the presence of spatio-temporal contiguity, any additional information seems to add little value. Alternatively, it is conceivable that the additional information confused the participants or led to processes that we did not anticipate. For example, purportedly supportive bottom-up input could be interpreted top-down as preventive information, thereby disrupting encoding. Yet, this consideration is purely speculative, and many other scenarios are conceivable. However, if one conclusion regarding bottom-up input can be drawn with certainty, it is that moderating bottom-up input is more challenging than we had anticipated.

Regarding top-down input, we used a p-curve analysis to further investigate the effect of our mixed results. We included (a) the findings of the experiments presented above (i.e., Experiments 1 & 2) and (b) the findings presented in Sperlich and Unkelbach (2022). Because non-statistical results are excluded from p-curve analyses, six values remained, $F(1,122) = 12.16$, $F(1, 136) = 25.12$, $F(1, 120) =$

11.09, *F*(1, 182) = 28.39, *F*(1, 146) = 5.65, and *F*(1, 59) = 8.08. The p-curve depicts a power estimate of 94%, which aligns with our two failures to find an effect (i.e., at least 1 out of 6 should probabilistically show no effect). As Figure 5.7B shows, the p-curve indicates evidential value for our findings. Thus, the analysis indicates that the top-down effect is likely to exist. However, an investigation of the influence of top-down input with decreased proximity of CS and US would still be interesting; a weaker bottom-up input could potentially leave even more room for the influence of top-down processes.

Figure 5.7B

A P-Curve Analysis for Top-Down Effects in Sperlich & Unkelbach (2022) and Experiments 1 and 2.



Note. The p-curve includes 6 statistically significant results (i.e., $p < .05$). Two additional results were entered but excluded from the p-curve and its analysis as they were $p > .05$. The blue line depicts the actual observed p-curve from the set of results. The red dotted line depicts how the line of p-values would look like if there would be no effect. The green line depicts how the p-curve would look like under 33% power. Created with p-curve app (4.06; https://www.p-curve.com/app4/).

5.8 Conclusion

While EC is deemed a robust effect in the laboratory (e.g., Walther, 2002), there is hardly any evidence for EC in the real world (De Houwer et al., 2001). Thus, Sperlich and Unkelbach (2022) proposed the dual-force model of EC in an attempt to explain this lack of evidence. They suggested that bottom-up and top-down input are required for EC to occur. We tested this assumption by implementing supporting top-down input (i.e., existing knowledge structures) on the organism's side. Our findings add to previous research, that could demonstrate the importance of top-down influences for EC. Yet, manipulating the bottom-up input did not yield significant results. This observation is not readily reconcilable with the dual-force model. However, our findings are consistent with the post-hoc explanation of spatio-temporal proximity representing such a strong bottom-up influence that any further bottom-up information adds little value to the encoding of CS and US into a shared episode.

Chapter 6: The Evaluative Information Ecology in Social Environments versus EC Experiments

In the previous chapter, four experiments were presented to test the dual-force perspective on EC. The top-down input's influence was shown in the laboratory; the experiments in which the bottom-up input was manipulated did, however, not yield significant results. One explanation might be that the close spatio-temporal contiguity of CS and US is such a strong manipulation that additional information is either ignored or leads to confusion. The striking relevance of stimulus contiguity has already been shown by some researchers (e.g., Ingendahl, Vogel et al., 2023; Jones et al., 2009; Hughes, Mattavelli et al., 2018). For example, Rydell and Jones (2009) presented participants with three stimuli in sequential order. First, a positive or negative US, then a CS, followed by another US with the opposite valence of the first US. In trials with partial overlap for US_1 (and no overlap for US_2), the CS took the valence of US_1 (i.e., the US with higher contiguity). In trials without overlap (i.e., when spatio-temporal proximity was held constant), the valence transfer took place between CS and negative US. Thus, this study is interesting for two reasons: (a) it shows that spatio-temporal contiguity plays a crucial role for the encoding of CS and US into the same mental episode and (b) it points to a negativity bias in EC. In the next section, this second point is examined in more detail. Even though Rydell and Jones (2009) found a negativity bias in their design, meta-analytical data provides evidence that EC experiments usually do not exhibit valence asymmetries. One explanation for this finding might be the artificial structure of experiments: the symmetrical distribution of positive and negative information does not correspond to the 'natural' distribution of the social environment as proposed by the EvIE model. Thus, the subsequent section explores whether variations in the frequency of positive and negative information align with the predictions of the EvIE model (Unkelbach et al., 2019), resulting in negativity and positivity biases.

This chapter is based on the following manuscript:

Sperlich, L. M., & Unkelbach, C. (2024). Why is there no negativity bias in Evaluative Conditioning? A cognitive-ecological approach. University of Cologne, Social Cognition Center Cologne. (manuscript under review at Journal of Personality and Social Psychology)

Please note that certain modifications were made to the headings, citation style, and formatting to align with the layout of this dissertation. No changes were made to the content of the article.

The re-analysis by analytic data was provided by Prof. Dr. Wilhelm Hofmann (Ruhr-University Bochum).

Why is There no Negativity Bias in Evaluative Conditioning? A Cognitive-Ecological Answer

Abstract

Evaluative Conditioning (EC) is the change of a conditioned stimulus (CS)'s evaluation due to its pairing with an unconditioned stimulus (US). While learning typically shows negativity biases, we found no such biases in a re-analysis of meta-analytic EC data. We provide and test a cognitive-ecological answer for this lack of a negativity bias. We assume that negativity effects follow from ecological differences in evaluative information's distributions (i.e., differential frequency). Accordingly, no negativity bias emerges because positive and negative information is equally frequent in most EC experiments. However, if negative (or positive) information is rare, we predict a negativity (positivity) bias. We tested this prediction in four pre-registered experiments (three laboratory-based, N = 394, one online, N = 192). As predicted, if negative USs were rare, a negativity bias followed. However, if positive USs were rare, we also observed positivity biases in participants' CS evaluations. These data support a cognitive-ecological explanation of valence asymmetries and partially explain why EC experiments show no negativity bias: Typical EC designs do not reflect the ecological information structure that contributes to a negativity bias in the first place.

Keywords: evaluative conditioning, attribute acquisition, negativity bias, information ecology, valence asymmetry

6.1 Why is there no negativity bias in Evaluative Conditioning?

Bad is stronger than good. This claim by Baumeister and colleagues (2001) holds for many domains in psychology; that is, negative information has more influence than positive information. For example, in social psychology, fewer instances are required to infer a disposition from unfavorable traits compared to favorable ones (Rothbart & Park, 1986), negative information outweighs positive information in impression formation (e.g., Amabile & Glazebrook, 1982; for a review, see Skowronski & Carlston, 1989), negative information is often more diagnostic than positive information (e.g., in the morality domain, Skowronski & Carlston, 1987; see also Fiske, 1980), negative words are classified faster than positive words (e.g., Nasrallah et al., 2009), and negative information leads to stronger EEG responses (Ito et al., 1998; see also Rozin & Royzman, 2001; Unkelbach et al., 2020, for further reviews on negativity biases). One may observe similar patterns in several other disciplines, such as behavioral decision-making (e.g., 'losses loom larger'; Tversky & Kahneman, 1991) or media psychology (e.g., 'if it bleeds it leads'; Robertson et al., 2023).

Given this evidence for the stronger impact of negative information over positive information, we expected a differential effect of positive and negative information in learning, specifically in the widely discussed paradigm of Evaluative Conditioning (EC; see Moran et al., 2023). EC is defined as the change in liking of a conditioned stimulus (CS) due to its pairing with an unconditioned stimulus (US) of positive or negative valence (De Houwer, 2007). EC is most relevant for explaining attitude acquisition (see Vogel & Wänke, 2016) but is also relevant in the areas mentioned above of stereotypes (e.g., French et al., 2013) and impression formation (e.g., Gast & Rothermund, 2011).

EC experiments should thereby constitute a prime candidate for the stronger impact of negative information. However, across many experiments in our lab, we found symmetrical effects of pairing positive and negative USs with CSs on CS evaluations. In addition, when we re-analyzed the data from the 2010 meta-analysis by Hofmann et al. (2010), we found no evidence of a negativity bias in EC (see below).

Given EC's theoretical and practical importance (see Moran et al., 2023), this apparent lack of a well-established phenomenon in evaluative learning (i.e., EC) deserves some consideration. More importantly, beyond the specific paradigm, answering this question also addresses the underlying processes of negativity advantages in social psychology. In the remainder, we first report a re-analysis of data from Hofmann et al.'s 2010 meta-analysis that supports our claim that there is no negativity bias in EC. We then provide two explanations of negativity bias (evolutionary vs. ecological) and argue that the cognitive-ecological approach might explain the absence of negativity bias in EC. We investigate our hypothesis in four pre-registered experiments, showing that we can create both a negativity bias and a positivity bias in EC as a function of the experiment's ecology. Finally, we will discuss the results considering evolutionary and cognitive-ecological explanations for negativity bias.

6.1.1 A Negativity Bias in EC – Meta-Analytic Evidence

As stated, we never observed a negativity bias in EC in our laboratory. To corroborate our personal experiences, we re-analyzed meta-analytic data from Hofmann and colleagues (2010). The meta-analysis contained 214 studies. We could only include studies that allowed us to compare the mean of positively (L = liked) and negatively paired CSs (D = disliked) with neutral CSs (N = neutral). We could not use studies without neutral CSs or with indirect measures, as the EC effect is inherently relative in these studies. To further maximize comparability, we only took studies using within-subject designs (i.e., "evaluations provided by the same participants", p. 399) that contained pre-post measures of CSs. We computed the influence of positive USs (i.e., neutral vs. liked: N-L) and negative USs (i.e., neutral vs. disliked, D-L) within participants. After the exclusion of outlier studies (i.e., studies that report effect sizes with more than 3 SDs from the mean effect size; Lipsey & Wilson, 2001; Nelson & Kennedy, 2009), we could analyze 56 studies that report N-L effect size contrasts, and 59 studies that report N-D effect size contrasts.

We used the meta-analytic ANOVA approach to test for the hypothesized negativity bias (see Hofmann et al., 2010, for details). We computed the variance in effect sizes that is "explained by the categorical variable (QB) as an indicator of variability between group means and the residual remaining portion (QW) as an indicator of variability within groups. […] QB is tested for significance against a chi-square distribution with df = j - 1 (where j is the number of categories or groups). A significant between-groups effect indicates that the variance in effect sizes is at least partially explained by the moderator variable” (p. 401).

Descriptively, in line with a negativity bias, we found that effect sizes for N-L contrasts were slightly smaller (d = .33, SE = .056) than effect sizes for N-D contrasts (d = .40, SE = .055). However, this difference was not statistically significant, QB(1) = 0.96, p = .327. When assuming a withincorrelation of .50 (cf. Borenstein et al., 2009), the calculated mean difference of ds was MDD = -0.022 (VDD = 0.0008), 95%CI [-0.078, 0.035]. Thus, the re-analysis confirmed our personal experiences that there is no negativity bias in EC. We next consider two explanations for negativity biases that might shed light on this absence of evidence for negativity bias in EC.

6.1.2 Evolutionary versus Cognitive-Ecological Explanations of Negativity Bias

One may explain negativity biases on two levels: A distal and a proximal level (Brunswik, 1955). For example, if one observes a negativity bias in impression formation (e.g., Fiske, 1980), it might be due to higher attention during learning and the resulting better memory for negative information. These cognitive processes provide a proximal explanation. However, they do not answer the distal question of why negative information receives more attention. We distinguish between two explanations on the distal level, an evolutionary (i.e., phylogenetic) explanation, and a cognitive-ecological (i.e., ontogenetic) explanation (see also Rozin & Royzman, 2001). To be sure, these are not mutually exclusive; similar to many robust psychological phenomena, more than one explanation most likely does apply. In addition, they are not exhaustive, although on the distal level, these two explanations cover most of the available theories (see also Unkelbach et al., 2019, 2020).

Evolutionary Explanation. The tenet of an evolutionary explanation, used here as an umbrella term for many similar theories, is that negativity bias(es) are adaptive (e.g., Taylor, 1991; Baumeister et al., 2001). Adaptiveness means that organisms showing a negativity bias are more likely to pass their genes to the next generation than organisms that do not. This explanation is easily illustrated with some simplifications for attention and memory, respectively. If an organism overlooks a predator, it will not pass its genes. Conversely, if the organism overlooks a mating opportunity, there might be other possibilities in the future. If an organism observes that certain berries are poisonous and others a nutritious, then it is adaptive to remember the poisonous berries because eating them prevents the passing of the genes. Conversely, not remembering the nutritious berries does not prevent other food possibilities in the future.

Relatedly, negative outcomes are also more extreme. However, several authors pointed out that a negativity bias due to greater extremity is of lesser interest, as it reduces the bias to an extremity bias (see Peeters & Czapinski, 1990; Unkelbach et al., 2019), and many negativity biases are found even when extremity is constant for positive and negative outcomes (e.g., Hilbig, 2009).

The evolutionary explanation nevertheless assumes that the negativity bias resides in the negative stimuli's valence without prior learning or situational influences. If organisms need to learn that attention to predators is advantageous, there would be no phylogenetic advantage. As Rozin and Royzman (2001, p. 314) stated: "[…] it is quite reasonable to suppose that the negative bias is a built-in predisposition. […] The opportunities for gradual learning to avoid death-threatening events may be minimal."

Cognitive-Ecological Explanation. The tenet of a cognitive ecological explanation is that negativity bias(es) result from basic cognitive processes that interact with the ecology (e.g., Fiske, 1980; Lewicki et al., 1992; Unkelbach et al., 2021). To illustrate, one might assume that cognition allocates more attention to infrequent events than frequent events due to ontogenetic learning. To explain negativity biases, one then needs the additional assumption that negative information is less frequent in the environment (see Rozin & Royzman, 2001; Unkelbach et al., 2019). Thus, the organism may attend to and remember negative information more because it is less frequent than positive information.

The cognitive-ecological explanation assumes that the ecology's structural properties in interaction with cognitive processes determine the bias. For example, in an ecology where negative information is frequent, the organism should attend more to positive information. In other words, in a world filled with predators and where most berries are poisonous, the one safe spot and the nutritious berries should receive more attention and better memory (see Alves et al., 2015, 2018; for empirical tests of such reversals).

6.1.3 Relating the Explanations to Negativity Bias in EC

An evolutionary explanation of negativity bias would predict differences in positive and negative US's influence on CS evaluations. Indeed, there is evidence for phylogenetical influences on learning, for example, under the label of 'preparedness' (Seligman, 1970; Seligman & Hager, 1972). However, the strongest negativity bias in learning from pairings is probably found for disgust (see Rozin & Haidt, 2013, for a review). As Rozin and Royzman (2001) noticed: "Brief contact with a cockroach will usually render a delicious meal inedible. The inverse phenomenon – rendering a pile of cockroaches on a platter edible by contact with one's favorite food – is unheard of" (p. 296). Assuming that many of the stimuli used in EC experiments (e.g., from the IAPS database; Alves & Imhoff, 2022; Houben et al., 2010; Pleyers et al., 2009) included disgusting stimuli, a negativity bias in EC would have been a viable hypothesis.

The cognitive-ecological explanation (Unkelbach et al., 2020) would predict a negativity bias in EC only if the ecology of an EC experiment reflects the structural properties of a standard ecology; most importantly, that negative information is less frequent than positive information. Regarding the higher frequency of positive information, it is apparent that most EC experiments do not reflect a standard ecology. Rather, for reasons of good experimental practice, positive and negative information is equally frequent in typical EC experiments. Such a symmetric setup does not allow for negativity bias due to differences in frequency without additional assumptions. For example, one would need to assume that people apply the environmentally learned asymmetric information distribution to the new ecology of the EC experiment. Without such strong additional assumptions, the cognitive-ecological explanation does not predict a negativity bias in EC.

Unkelbach et al. (2019, 2020) also proposed another structural property of a standard ecology that may lead to negativity biases: negative information is more diverse than positive information. However, the case for a higher diversity of negative information is more complex and similarity as an explanatory construct is more difficult to assess and depends on context factors (see Koch et al., 2022; Medin et al., 1993). For the present case, we will focus on the differential frequency of positive and negative information.

Thus, given the substantial effects of negative stimuli on learning in other domains (e.g., disgust from pairings, Rozin & Haidt, 2013), the lack of a negativity bias in EC might seem surprising. However, if one assumes that negativity bias arises not only due to phylogenetic pressures but also due to the structural properties of the respective ecology, the symmetrical evaluative ecology of most EC experiments might prevent a negativity bias.

The cognitive-ecological explanation allows testable predictions for the absence of negativity bias, namely that if an EC experiment's ecology reflects the standard ecology, then a negativity bias should follow. Further, if an EC experiment's ecology reverses the structural properties, then a positivity bias should follow (see Alves et al., 2017b; Walasek & Stewart, 2015). The following four experiments test this possibility. We manipulated the frequency distribution of positive and negative information (i.e., frequent vs. rare).

The varying frequency may translate into negativity or positivity biases (i.e., stronger influences of rare information) via several interrelated and overlapping well-established cognitive and judgment processes. Prominently, rare stimuli are more diagnostic (e.g., Fiske, 1980; Skowronski & Carlston, 1989) and have therefore an attention and memory advantage. Similarly, infrequent stimuli are distinct and benefit therefore from an isolation effect (von Restorff, 1933), that is, a memory advantage for distinct stimuli. Relatedly, one may consider rare stimuli as the figure that is more visible compared to the background (see Wagemans et al., 2012, for an overview). Finally, rare stimuli may be comparatively contrasted away from the frequent stimuli in judgments (Bless & Schwarz, 2010; see Unkelbach et al., 2019, for an overview). The following experiments do not test these proximal cognitive processes but the proposed ecological explanation.

6.2 Preview of Experiments

The four presented experiments used a simultaneous conditioning procedure, showing CS and US next to each other (e.g., Stahl & Heycke, 2016). Unlike typical EC experiments, we varied the frequency of positive and negative USs, thereby constructing either a standard or reversed ecology. In a standard positive ecology, positive USs were frequent and negative USs rare. In a reversed negative ecology, negative USs were frequent and positive USs were rare. Given this setup, a cognitive-ecological explanation would predict a main effect of the ecology.

Let us clarify this predicted pre-registered ecology main effect. In the positive ecology, we assume that participants rate CS_{pos} (i.e., CSs paired with positive USs) in the positive ecology less positively than CS_{pos} in the negative ecology, while they rate CS_{neg} (i.e., CSs paired with negative USs) more negatively. Thus, ratings in the positive ecology become overall less positive (i.e., strong CS_{neg} effects, weak CS_{pos} effects), while this pattern should reverse in the negative ecology (i.e., weak CS_{neg} effects, strong CS_{pos} effects).

Another option would have been a comparison with a neutral CS in each ecology. However, we decided against this possibility for two reasons. First, people might make inferences about neutral CSs in positive and negative ecologies (e.g., without a basis for evaluation, they might use the US base-rate valence). Second, people who observe more positive outcomes might be in a better mood than people who observe more negative outcomes and judge the neutral CS accordingly. Thus, we opted for comparing CSs paired with USs across different ecologies. We must therefore assume that participants use the scale similarly across conditions. However, this is a standard assumption in between-participant designs.

Experiment 1a used the outcomes of a lottery as the USs, which is a natural setting for manipulating the frequency of positive and negative outcomes (i.e., lotteries with frequent or infrequent wins and losses). Experiment 1b replicated Experiment 1a with a slight modification to address the potential ambivalence of the negative outcomes (i.e., 'not winning' vs. 'losing'). Experiment 2 compared the lottery setting with a typical EC picture-picture condition. Finally, Experiment 3 replicated Experiment 2 but eliminated the potential confound that we compare averaged ratings for one valence (i.e., related to the frequent USs) with a single rating of the other valence (i.e., related to the rare USs).

6.3 Transparency and Open Science

We pre-registered predictions and sample sizes for all four experiments. We did not know the expectable effect size of differential US frequencies. We typically find EC effects within-participants with samples of N = 25, which aligns with the meta-analytic effect size reported by Hofmann et al. (2010). As we aimed to moderate the CS ratings between conditions, we quadrupled the sample (see Simonsohn, 2014). To save laboratory resources, we pre-registered a sequential analysis with 200 participants (400 in Experiment 2; details see below) based on the recommendations by Lakens (2014) for Experiment 1a and kept this strategy for the following three Experiments. A post-hoc *p*-curve analysis (see Simonsohn et al., 2015) of the relevant main effects' *F*-values places our average power at 79%.

We report all data exclusions, which were all pre-registered, all manipulations, and all measures. The data, pre-registrations, analysis code and materials are available from the Open Science Framework (https://osf.io/9qbhr/?view_only=0167ad94e3a7437ca3cfab48d9358b3f).We conducted this research according to the American Psychological Association's (2017) Ethical Principles in the Conduct of Research with Human Participants.

We used the R software (R Core Team, 2021) to analyze our data. The following R packages were used: afex (Singman et al., 2016), dplyr (Wickham et al., 2021), ggpubr (Kassambara, 2020), magrittr (Back & Wickham, 2020), plyr (Wickham, 2011), psych (Revelle, 2021), reshape2 (Wickham, 2007), readr (Wickham & Hester, 2021), and stringr (Wickham, 2019).

Due to our pre-registered exclusion criteria, Experiments 1a and 1b excluded a substantial proportion of participants (i.e., 33% and 30%, respectively). However, we replicate the relevant main effects in all four experiments.

In addition to the four experiments reported here, we conducted two further experiments where we manipulated the frequency of positive and negative USs. The supplements report the data from these two experiments. Chronologically, they represented the third and fourth experiments in the series. These two experiments showed non-significant results (i.e., p = .051, and p = .257, for the ecology main effect, with N = 200 and N = 200, respectively). These experiments differed from the reported experiments in two relevant ways: First, they had fewer learning trials (i.e., 20 and 30 learning trials, respectively). Second, they employed a within-participants design rather than the between-participants of the present experiments. In addition, we conducted two experiments in which we attempted to manipulate the diversity of positive and negative USs. However, we failed to establish the differential diversity.

6.4 Experiment 1a: You Won versus You Did Not Win

In Experiment 1a, participants played a wheel-of-fortune lottery. We manipulated the frequency of positive and negative outcomes to create a positive or negative ecology. The lottery outcomes served as USs ('you won' vs. 'you did not win'). Simple geometric shapes with different colors served as CSs. We expected that rare outcomes should have a stronger influence on participants' CS evaluations compared to frequent outcomes, independent of valence.

6.4.1 Method

Participants and Sample Size. We pre-registered a sequential analysis based on the recommendations by Lakens (2014). We planned to gather data from 200 participants but analyze the data after collecting half the data with α = .0294 for the predicted main effect of ecologies. We had

102 participants in the first phase. We observed the predicted effect and stopped collecting data (see Lakens, 2014).

We collected the data laboratory-based on a city-based university campus. Research assistants approached people and invited them to participate for a payment of 2 Euro (about $2.30). The sample included students and pedestrians. Two participants could not finish the experiment due to technical problems, and thus we recruited two additional participants, resulting in 102 participants.

Materials. We programmed the experiment with OpenSesame (Mathôt et al., 2012). A wheel-of-fortune on the screen represented the lottery. As stated, the information: 'You won' (positive valence) and 'You did not win' (negative valence) served as USs, and simple geometric shapes with different colors served as CSs (from Förderer & Unkelbach, 2011). The program randomly selected five CSs from a pool of 12 CSs for each participant.

Design. We varied US valence within participants, resulting in a 2 (positive ecology vs. negative ecology; between) x 2 (US valence: positive vs. negative, within) mixed design.

Procedure. Upon arrival, experimenters accompanied participants to individual PC workstations and started the OpenSesame program. After providing informed consent, experimenters started the computer program. The computer program assigned participants alternatingly to either the 'positive ecology' condition or the 'negative ecology' condition and instructed them on how to spin the wheel-of-fortune. To ensure outcome valence, participants in both ecology conditions read that they would spin the wheel to collect points, with each point being equivalent to the value of 0.01€. They would receive the accumulated sum as a bonus at the end of the experiment. To keep payment constant, experimenters always rounded up the points (96 points or 24 points, depending on the ecology condition) to 100 points, resulting in a 1€ bonus for everybody.

After instructions, participants did spin the wheel at the beginning of each round by clicking on it. After spinning the wheel, a blank screen appeared for 500ms, followed by the presentation of the outcome ('you won' vs. 'you did not win') and the respective CS form, which constituted the CSUS pairing (2000ms). The CS forms always appeared on the screen's left half, and the outcome US always appeared on the screen's right half. After 2000 ms, another blank screen appeared for 500 ms before the next lottery round. Participants played 120 lottery rounds.

Depending on the ecology condition, participants won 96 times and did not win 24 times, or won 24 times and did not win 96 times. In the positive ecology condition, the program randomly assigned four CSs to the frequent positive US (i.e., 'You won') and one CS to the rare negative US (i.e., 'You did not win'). In the negative ecology condition, the program assigned one CS to the rare positive US (i.e., 'You won') and four CSs to the frequent negative US (i.e., 'You did not win'). The program paired each CS 24 times with its respective positive or negative outcome, and each CS appeared once before a given CS could appear the next time. In this sequence of five CSs, the rare US and its respective CS appeared after three frequent US outcomes and their respective CSs to highlight that it is a rare outcome (cf. Schmidt & Schmidt, 2017).

After the 120 lottery rounds, the CS rating phase started. The program instructed participants to rate their liking of different forms on a 5-point-scale ranging from -2 ('not at all') to +2 ('very much') with 0 ('neutral') as the center of the scale. Each form appeared once. The program randomized the CS rating order for each participant. After the rating phase, participants provided demographic information, reported their hypotheses about the study in a free format, and reported their concentration during the experiment. At the end, the experimenters debriefed and paid them.

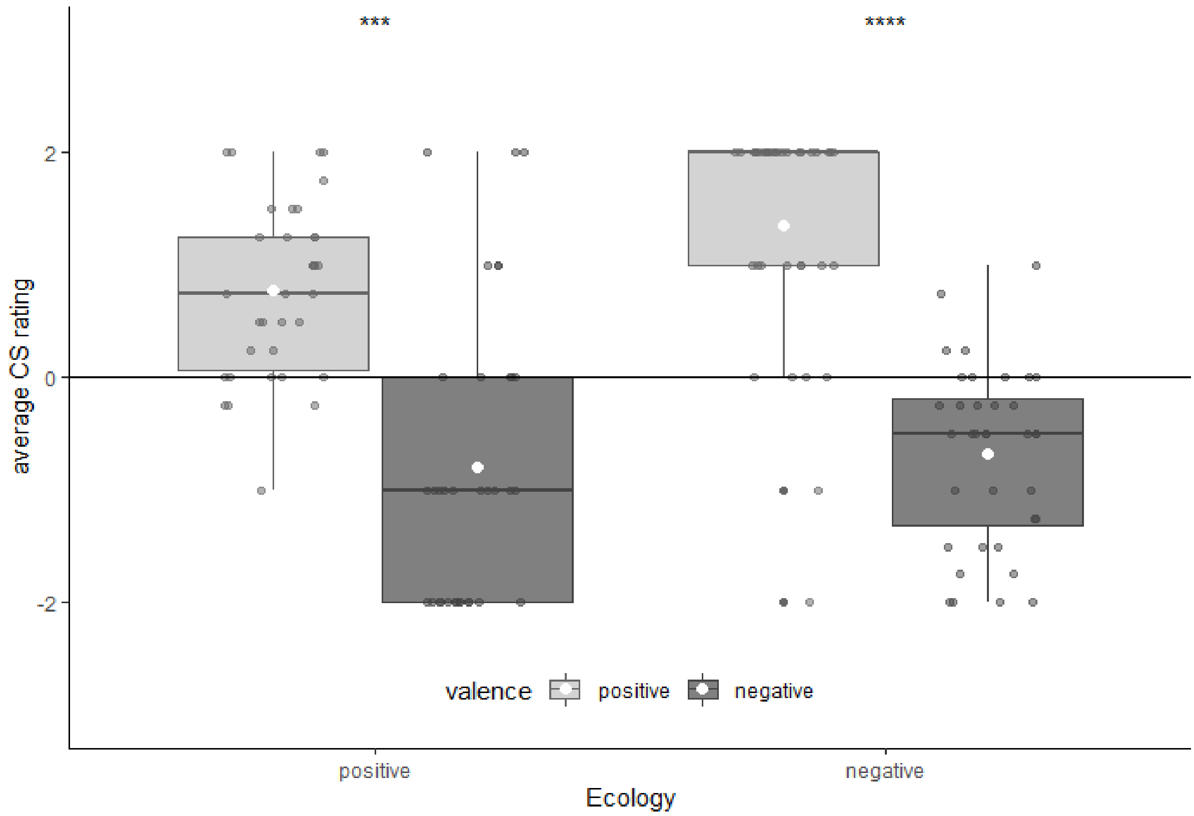
6.4.2 Results

Seventy participants reported potential hypotheses. Two research assistants coded these free responses. None of these 70 participants reported hypotheses that aligned with the experiment's hypothesis. However, as pre-registered, we excluded 29 participants because they reported low concentration. In addition, we excluded one participant due to their age (i.e., they could not legally provide consent). Thus, we retained 70 participants in the sample (42 women, 24 men, 1 prefer not to say; M_{age} = 24.91, SD_{age} = 7.17).[26] Thirty-four were in the positive and 36 in the negative ecology conditions.

CS ratings. Figure 6.4 presents participants' CS ratings averaged within US valence in the positive ecology and the negative ecology conditions. We analyzed these data with an ecology (positive ecology vs. negative ecology) by US valence (positive vs. negative) mixed ANOVA with repeated measures on the latter factor.

Figure 6.4

Combined Box and Scatter Plots of Participants' CS Ratings in Experiment 1a as a Function of US Valence (Positive vs. Negative; Within Participants) and Ecology (Positive vs. Negative; Between Participants).



Note. Higher values indicate more positive CS evaluations. The box plots constitute the first quartile, third quartile, and the median. The circle within a box plot indicates the mean. Each dot represents a CS evaluation.

[26] Due to technical problems, demographical data from three participants were not available.

This ANOVA showed a valence main effect, $F(1, 68) = 87.89$, $p < .001$, $\eta_p^2 = .56$, which constituted a typical EC effect. Participants' liking ratings of CSs presented with positive USs were on average higher ($M = 1.08$, $SD = 0.93$) than their liking ratings of CSs presented with negative USs ($M = -0.74$, $SD = 1.06$).

More importantly, however, we found the predicted main effect for ecology, $F(1, 68) = 6.68$, $p = .012$, $\eta_p^2 = .09$.[27] Participants rated a CS paired with a rare positive US more positive (i.e., in the negative ecology, $M = 1.36$, $SD = 0.99$) compared to CSs paired with frequent positive USs (i.e., in the positive ecology; $M = 0.77$, $SD = 0.77$). Vice versa, participants rated a CS paired with a rare negative US more negative (i.e., in the positive ecology; $M = -0.79$, $SD = 1.30$) compared to CSs paired with frequent negative USs (i.e., in the negative ecology; $M = -0.68$, $SD = 0.80$). This joint increase results in the overall main effect that participants rated CSs less positive in the positive ecology ($M = -0.01$, $SD = 1.32$) than CSs in the negative ecology ($M = 0.34$, $SD = 1.36$).

6.4.3 Discussion

Experiment 1a showed a typical EC effect. Participants liked forms more that appeared with positive outcomes compared to forms that appeared with negative outcomes. More critically, it showed the predicted main effect for ecology. Participants rated CSs presented with rare USs more extreme than CSs presented with frequent USs, resulting in overall more positive CS evaluations in the negative ecology and more negative CS evaluations in the positive ecology. Thus, in line with a cognitive-ecological explanation, rare US outcomes more strongly impacted CS evaluations compared to frequent US outcomes.

In typical EC experiments, positive and negative USs appear with equal frequency. Researchers employ equal frequencies with good reason; for example, to avoid mood effects. In Experiment 1a, one might speculate that frequent positive outcomes lead to a better mood than

---

[27] Exclusions did not change the main effect for valence, $F(1, 98) = 92.73.46$, $p < .001$, $\eta p^2 = .49$. However, without exclusions, the ecology main effect was not significant, $F(1, 98) = 3.92$, $p = .051$, $\eta p^2 = .04$, but there was a significant interaction effect, $F(1, 98) = 5.65$, $p = .019$, $\eta p^2 = .06$.

frequent negative outcomes, which might influence CS ratings, as in a mood-as-information effect (e.g., Schwarz & Clore, 2003). However, the present critical main effect is the opposite of a mood effect.

Of relevance, the equal frequency of positive and negative outcomes does not represent people's standard ecology in which negativity biases appear. In standard ecologies, positive outcomes are more frequent than negative outcomes (see Rozin & Royzman, 2001), and most likely, the ratio goes beyond 4:1 (e.g., the ratio of hugs to slaps, the ratio of contracts honored to broken, the ratio of smiles to frowns; see Unkelbach et al., 2019, for a review).

If negative outcomes were rare in the positive ecology condition, these had a stronger influence on CS evaluations, implying a negativity bias. Notably, we observed this stronger influence also for positive outcomes. Thus, Experiment 1a provides initial evidence for the influence of frequency on the evaluation of stimuli in EC.

However, Experiment 1a also raised some questions. First, we excluded about 30% of the participants. While we pre-registered the exclusion criteria, the remaining sample is comparatively small. In addition, the information 'You did not win' still contains the positive term 'win'. Previous research showed that ambivalence can be evaluatively conditioned (Glaser et al., 2018). The US 'You did not win' might represent an ambivalent stimulus rather than a clearly negative US, and we arguably did not establish two clear-cut evaluative ecologies. The ecology effect might hinge on the ambivalent nature of the negative USs. We thus aimed to replicate Experiment 1a with a stronger manipulation of negative US valence.

6.5 Experiment 1b: You Won versus You Lost

The main change from Experiment 1b to 1a was that the lottery outcome stated 'You lost', leaving no room for ambiguity regarding the outcome's negativity.

6.5.1 Method

Participants and sample size. As in Experiment 1a, we pre-registered a sequential analysis (cf. Lakens, 2014) with a planned sample size of N = 200 and one intermediated analysis at n = 100 with α = .294. As we did not find the predicted effect with n = 100, we moved on to collect the full pre-registered sample. We collected the data in the same laboratory as Experiment 1a, and research assistants checked whether participants had already participated in Experiment 1a. A sample of 201 people, students, and pedestrians, participated for a payment of 2 Euro (about $2.30).

Materials, design, and procedure. Everything was similar to Experiment 1a, but we changed the negative lottery outcome from 'You did not win' to 'You lost'.

6.5.2 Results

Six participants did not finish the study. In addition, as pre-registered, we excluded 60 participants because they reported low concentration. Thus, we retained 135 participants in the sample (87 women, 47 men; M_{age} = 23.81, SD_{age} = 9.35).[28] Seventy were in the positive ecology condition, and 65 were in the negative ecology condition. Of these 135 participants, as in Experiment 1a, none reported suspicions about the experiment's hypothesis.

CS ratings. Figure 6.5 presents participants' CS ratings averaged within US valence in the positive and negative ecology. We analyzed these data with an ecology (positive ecology vs. negative ecology) by US valence (positive vs. negative) mixed ANOVA with repeated measures on the latter factor.

We again found a typical EC effect, $F(1, 133) = 137.38$, $p < .001$, $\eta_p^2 = .51$; participants liked CSs presented with positive USs on average more (M = 0.82, SD = 0.94) than CSs presented with negative USs (M = -0.77, SD = 0.98).

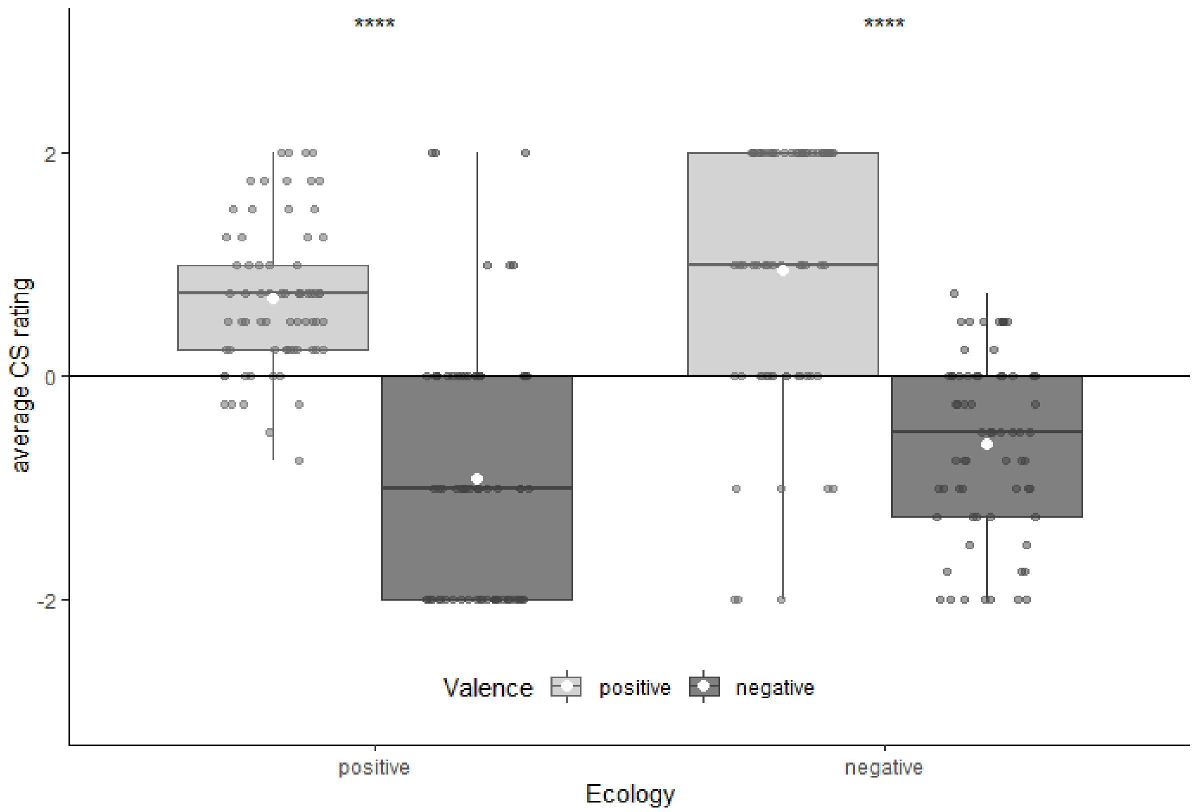
More importantly, we again found the expected main effect for ecology, $F(1, 133) = 8.99$, p = .003, $\eta_p^2 = .06$.[29] Participants rated a CS paired with a rare positive US more positive (i.e., in the negative ecology, M = 0.95, SD = 1.16) compared to CSs paired with frequent positive USs (i.e., in the positive ecology; M = 0.70, SD = 0.65). Vice versa, participants rated a CS paired with a rare negative US more negative (i.e., in the positive ecology; M = -0.91, SD = 1.11) compared to CSs paired with

[28] Due to technical problems, demographical data from one participant were not available.
[29] Exclusion of participants did not change the effects. The valence main effect, $F(1, 193) = 182.16$, $p < .001$, ηp^2 = .49, as well as the ecology main effect, $F(1, 193) = 9.19$, $p = .003$, $\eta_p^2 = .05$, remained significant without exclusions.

frequent negative USs (i.e., in the negative ecology; *M* = -0.61, *SD* = 0.80). This joint increase results in the overall main effect that participants rated CSs less positive in the positive ecology (*M* = -0.11, *SD* = 1.22) than CSs in the negative ecology (*M* = 0.17, *SD* = 1.27).

Figure 6.5

Combined Box and Scatter Plots of Participants' CS Ratings in Experiment 1b as a Function of US Valence (Positive vs. Negative; Within Participants) and Ecology (Positive vs. Negative; Between Participants).



Note. Higher values indicate more positive CS evaluations. The box plots constitute the first quartile, third quartile, and the median. The circle within a box plot indicates the mean. Each dot represents a CS evaluation.

6.5.3 Discussion

Experiment 1b replicated Experiment 1a: The rare US outcomes again had a stronger effect on participants' CS evaluations compared to the frequent US outcomes, visible as an ecology main effect. Experiment 1b also clarified that this ecology effect is not due to the potentially ambivalent nature of the negative US. Again, in line with a cognitive-ecological explanation, we observed a negativity bias when the experiment resembled a standard ecology regarding its evaluative frequency distribution (i.e., frequent positive outcomes relative to rare negative outcomes). We observed a positivity bias if the ecology reversed (i.e., rare positive outcomes relative to frequent negative outcomes).

However, while our lottery setup lends itself nicely to manipulating the frequency of positive and negative outcomes, it is not a standard EC paradigm. Rather, the picture-picture paradigm, introduced by Levey and Martin (1975), is the prototypical EC paradigm (Dawson et al., 2007). A substantial number of EC experiments pair positively or negatively evaluated pictures as USs with neutral pictures (e.g., Baeyens et al., 1992; De Houwer et al., 2000; Gast & Rothermund, 2011; Schienle et al., 2001; Walther & Nagengast, 2006). Experiment 2 tested whether we may create positivity and negativity biases in the picture-picture paradigm.

6.6 Experiment 2: Picture-Picture Paradigm versus Lotteries

In Experiment 2, we investigated whether we may replicate Experiment 1a and 1b's findings using pictures instead of lottery outcomes. In addition, we also replicated Experiment 1b as a comparison standard.

6.6.1 Method

Participants. The COVID-19 pandemic forced us to close our laboratory. We thus conducted Experiment 2 online on Prolific. As we replicated Experiment 1b in one condition and used the picture-picture paradigm in the other, we pre-registered 400 participants (i.e., 200 per condition) in a sequential analysis with one intermediate analysis at n = 200 with α = .294. We did not expect an interaction. We stopped collecting data after the first phase with n = 200 participants on Prolific (after exclusion[30]) who received a payment of £1.13 (about $1.48).

[30] Data from 204 participants were collected. 4 participants terminated the study early, and were excluded by Prolific (i.e., Prolific does not count unfinished data as collected data; as preregistered, we excluded these data sets).

Materials. We programmed the experiment using Qualtrics to collect the data online. The lottery condition replicated Experiment 1b online. We used the same CSs as in the lottery condition for the picture-picture condition. As USs, we presented pictures from Förderer & Unkelbach (2012) showing subjectively cute or ugly animals (i.e., based on other participants' pre-ratings). To exclude the possibility of valence asymmetries due to differences in US appearance, we exclusively used animals with fur as positive and negative USs (e.g., dogs, cats, horses, etc.). The stimuli were pretested for valence, resulting in M = 61.77 (SD = 12.25) for positive USs and M = -46.72 (SD = 16.44) for negative USs on a scale from 0 to 100, with higher values indicating higher liking (for details, see supplement).

Design. Our design had four between-participants conditions resulting from the orthogonal combination of ecology (i.e., positive vs. negative) and US type (i.e., lottery outcomes vs. animal pictures). We varied US valence (i.e., positive vs. negative) within participants.

Procedure. After providing informed consent, the program randomly assigned participants to one of the four experimental conditions. The program instructed participants in the picture-picture condition to watch a slideshow and participants in the lottery condition to spin the wheel-of-fortune. To reduce the experiment's length for the online settings, we reduced the number of learning trials from 120 to 60.

The CS-US pairing phase in the lottery condition was similar to Experiment 1b. In the picture-picture condition, the survey randomly assigned the five CS forms to one US picture each (i.e., 'one-to-one' assignment; see Stahl & Unkelbach, 2009). In the positive ecology condition, the program randomly assigned four CS forms to the positive US pictures and one CS form to a negative US picture. In the negative ecology, the program randomly assigned four CS forms to negative US pictures and one CS form to a positive US picture. The program presented each CS-US pairing 12 times.

Participants did not collect points in the lottery condition to keep the US type conditions comparable. As in Experiments 1a and 1b, participants in the lottery condition had to click the wheel at the beginning of each round. Participants in the picture-picture conditions had to press a button to see the CS-US pairing. After this action, a blank screen appeared for 500 ms, followed by the presentation of the CS-US pairing (2000 ms) and another blank screen (500 ms). As in Experiment 1a and 1b, the program organized the CS-US pairings in sequences of five pairings, with the rare US at the fourth position (Schmidt & Schmidt, 2017).

After the 60 conditioning trials, the CS rating phase started, identical to Experiments 1a and 1b. After the rating phase, participants provided demographic information, and the Qualtrics survey provided them with a code for payment via Prolific.

6.6.2 Results

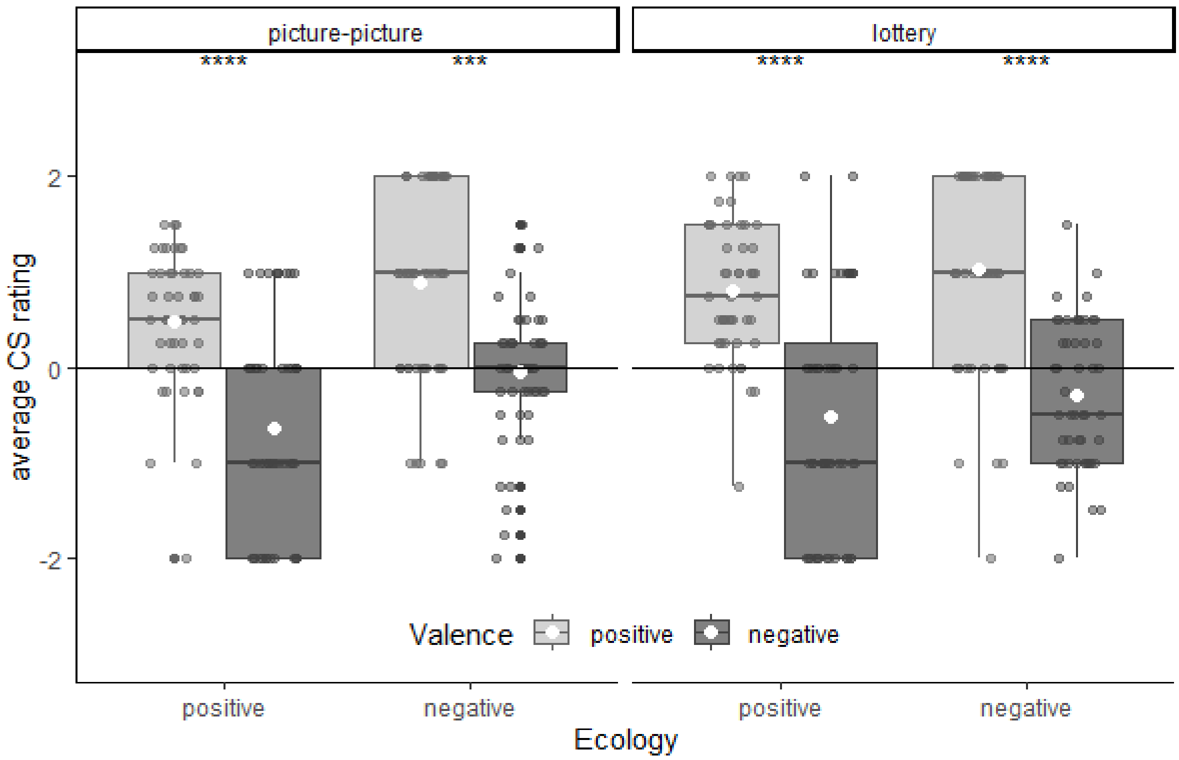
As stated, we pre-registered a sequential analysis (cf. Lakens, 2014) with a planned N of 400. Thus, we analyzed the data after collecting data from 200 participants with α = .0294 for the predicted difference between the two ecologies. As we observed this predicted effect, we stopped collecting data at this point. We increased the number of participants by 100% due to the increased number of conditions.

As pre-registered, we excluded 8 participants that indicated low concentration. Thus, we retained data from 192 participants for analysis (151 women, 40 men, 1 diverse; M_{age} = 40.09, SD_{age} = 13.11). Ninety-five were in the positive ecology condition, and 97 were in the negative ecology condition.

CS ratings. Figure 6.6 presents participants' averaged CS ratings within US valence as a function of ecology (i.e., positive vs. negative) within the US type conditions (i.e., lottery outcomes vs. pictures). We analyzed these data with a condition by ecology by US valence mixed ANOVA, with repeated measures on the latter factor.

Figure 6.6

Combined Box and Scatter Plots of Participants' CS Ratings in Experiment 1a as a Function of US Valence (Positive vs. Negative; Within Participants), Ecology (Positive vs. Negative; Between Participants), and US Type (Picture-Picture vs. Lottery Outcome; Between Participants).



Note. Higher values indicate more positive CS evaluations. The box plots constitute the first quartile, third quartile, and the median. The circle within a box plot indicates the mean. Each dot represents a CS evaluation. Observations outside the whiskers are outliers with more than 150% interquartile range distance from the next quartile.

The analysis showed the typical US valence main effect, $F(1, 188) = 125.06$, $p < .001$, $\eta_p^2 = .40$. Participants liked CSs presented with positive USs on average more ($M = 0.81$, $SD = 0.90$) than CSs presented with negative USs ($M = -0.37$, $SD = 0.99$).

Importantly, we again found the expected main effect for ecology, $F(1, 188) = 19.56$, $p < .001$, $\eta_p^2 = .10$.[31] Participants rated a CS paired with a rare positive US more positive (i.e., in the negative ecology, $M = 0.97$, $SD = 1.03$) compared to CSs paired with frequent positive USs (i.e., in the positive ecology; $M = 0.65$, $SD = 0.73$). Vice versa, participants rated a CS paired with a rare negative US more negative (i.e., in the positive ecology; $M = -0.58$, $SD = 1.15$) compared to CSs paired with frequent negative USs (i.e., in the negative ecology; $M = -0.16$, $SD = 0.75$). Again, this joint increase results in the overall main effect that participants rated CSs less positive in the positive ecology ($M = 0.04$, $SD = 1.14$) than CSs in the negative ecology ($M = 0.40$, $SD = 1.06$).

There was no significant main effect for US type, $F(1, 188) = 1.13$, $p = .290$, $\eta p^2 = .01$. Participants' CS ratings in the picture-picture condition did not differ significantly from participants' ratings in the lottery condition, and none of the interactions was significant, largest $F(1, 188) = 2.55$, $p = .112$, $\eta p^2 = .01$.

6.6.3 Discussion

Experiment 2 replicated the results of Experiment 1a and 1b in a typical picture-picture EC paradigm (Levey & Martin, 1975). We found the expected stronger influence of rare USs also when using pictures as USs. Experiment 2 thereby shows that one may create both negativity and positivity biases as a function of the experimentally created ecology. We thus tentatively conclude that one reason why there are no negativity biases in EC is that the evaluative information ecology of a typical EC experiment does not reflect the standard ecology, in which negative outcomes or stimuli are rare and positive outcomes or stimuli are frequent (see Unkelbach et al., 2019). However, we have one confound so far: Participants provided only one rating for a CS paired with the rare US, but several (i.e., four) ratings for CSs paired with frequent USs. This confound may contribute to our observed positivity and negativity biases. Several ratings may lead to an averaging effect; participants may rate some CSs extremely and others less extremely. Given successful conditioning, the single CS paired

[31] Not excluding participants with low concentration did not change these results. The main effect for valence, $F(1, 196) = 129.98$, $p < .001$, $\eta p^2 = .40$, as well as for ecology, $F(1, 196) = 20.81$, $p < .001$, $\eta p^2 = .10$, remained significant.

with the rare US may receive more extreme ratings, which would reduce our bias to a rating effect, while we assume an effect on the level of the CS's mental representation (i.e., the CS paired with the rare US should be indeed liked less/more). While one may also assume that averaging should also happen across participants for a single rating, it seems necessary to address this confound.

6.7 Experiment 3: Sampling Frequent Stimuli

With Experiment 3, we aimed to replicate the findings of Experiment 2 but address the rating confound. We implemented three changes. First, we dropped the lottery condition, using only Experiment 2's picture-picture condition. Second, the program randomly sampled only one of the CSs paired with the frequent USs for the CS rating phase, leaving participants with only two CSs to rate (i.e., one per valence). Third, we added a memory task to investigate whether memory is important for the occurrence of valence asymmetries. To spoil the outcome for the memory measure, memory was at ceiling the way we measured it in both conditions, and we did not find any differences on this measure.

6.7.1 Method

We conducted Experiment 3 after the end of the COVID-19 restrictions, and we could again collect data in our university labs. Experiment 3 was, therefore, again laboratory-based.

Participants and sample size. We again pre-registered a sequential analysis (cf. Lakens, 2014) with 200 participants. We collected data from 100 participants and then ran the planned analyses with $\alpha = .0294$ for the predicted ecology main effect. As we did not observe the predicted effect, we continued the data collection (cf. Lakens, 2014) and ended with a sample of 202 participants from our university campus. The sample was comparable to Experiments 1a and 1b. Participants received a payment of 2€ (about $2.30).

Materials. We programmed the experiment in OpenSesame, using the same CSs as in Experiment 2. As USs we used pictures of positive and negative dogs (for sources, see supplement).

Design. The computer program assigned participants alternatingly to the positive ecology or negative ecology conditions. We varied US valence within participants, resulting in an ecology (positive ecology vs. negative ecology; between) x US valence (positive vs. negative, within) mixed design.

Procedure. Upon arrival, experimenters accompanied participants to individual PC workstations and started the OpenSesame program. After providing informed consent, the program randomly assigned participants to one of the ecology conditions. In the conditioning phase, the program randomly sampled dog USs from 12 positively and 12 negatively US pictures. Everything else was similar to Experiment 2. Before the rating phase started, the program randomly sampled one of the frequently presented CSs. Thus, different from before, participants only rated two instead of five CSs. This change avoided the previous rating confound. After the rating phase, participants performed a memory test for US valence and US identity. Each participant judged the previously drawn CS paired with a frequent US valence and the CS paired with a rare US valence in both tasks. In the valence memory task, participants indicated whether the CS was presented with a positive or negative picture of an animal. As a third option, they could indicate that they did not know the answer. In the US identity task, the program presented the US that was presented with the CS in the conditioning phase (i.e., the correct US) and four distractor stimuli randomly drawn from the pool of USs not used for the respective participant. The order of USs was randomized for each participant. Participants were then asked to choose the correct US. After providing demographic data, they were debriefed and paid.

6.7.2 Results

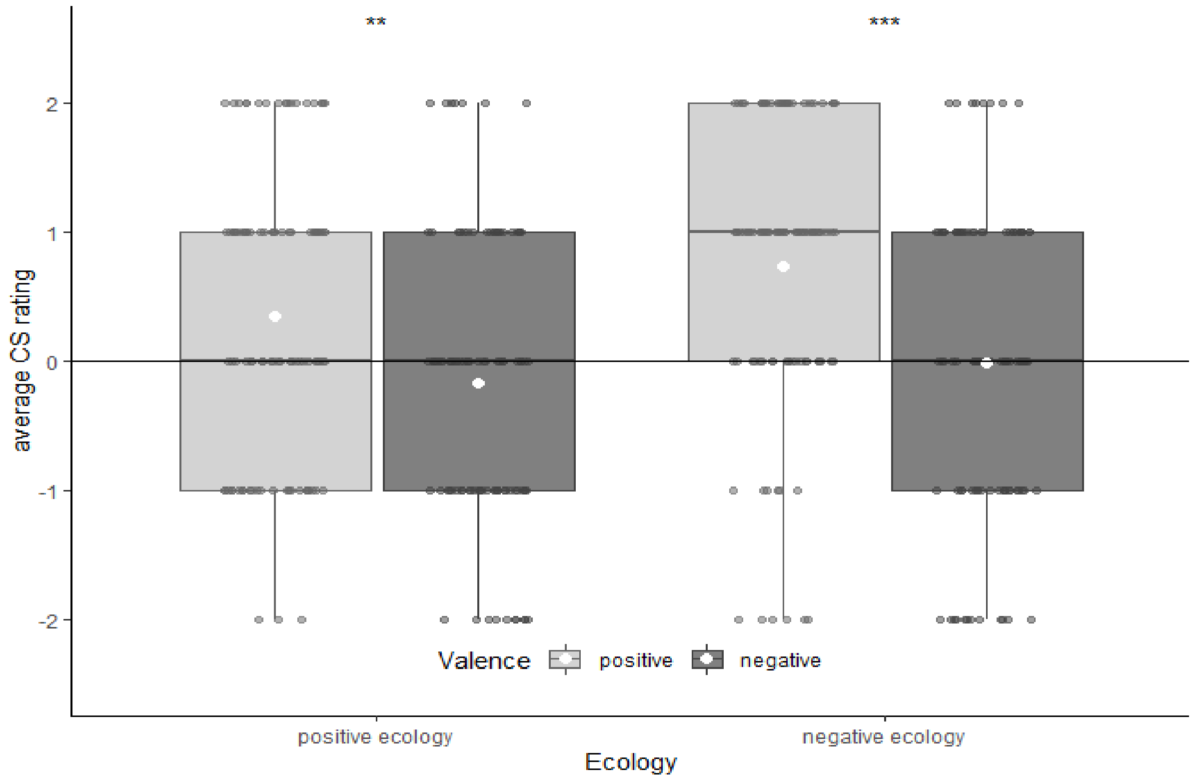
As pre-registered, we excluded nine participants who did not finish the study, one participant who indicated a low concentration, and two who could not provide consent due to their age. Additionally, one participant had missing values in the data set. Thus, we retained data from 189 participants for analysis (122 women, 67 men; M_{age} = 24.31, SD_{age} = 8.18).

CS ratings. Figure 6.7 presents participants' CS ratings averaged within US valence in the two ecology conditions. We analyzed these data with an ecology (positive vs. negative) by US valence (positive vs. negative) mixed ANOVA with repeated measures on the latter factor.

Again, there was a main effect for valence, $F(1, 187) = 26.17$, $p < .001$, $\eta_p^2 = .12$.[32] Participants liked CSs presented with positive USs on average more ($M = 0.54$, $SD = 1.14$) than CSs presented with negative USs ($M = -0.09$, $SD = 1.13$).

Figure 6.7

Combined Box and Scatter Plots of Participants' CS Ratings in Experiment 3 as a Function of US Valence (Positive vs. Negative; Within Participants) and Ecology (Slideshow: Positive vs. Slideshow: Negative; Between Participants).



Note. Higher values indicate more positive CS evaluations. The box plots constitute the first quartile, third quartile, and the median. The circle within a box plot indicates the mean. Each dot represents a CS evaluation.

[32] The effects did not change when including the participant with low concentration. The valence main effect, $F(1, 188) = 26.65$, $p < .001$, $\eta_p^2 = .12$ as well as the ecology main effect, $F(1, 188) = 6.22$, $p = .013$, $\eta_p^2 = .03$, remained significant.

As predicted, we also found the ecology main effect, $F(1, 187) = 6.37$, $p = .012$, $\eta_p^2 = .03$. Participants rated a CS paired with a rare positive US more positive (i.e., in the negative ecology, M = 0.74, SD = 1.12) compared to CSs paired with positive frequent USs (i.e., in the positive ecology; M = 0.36, SD = 1.13). Vice versa, participants rated a CS paired with a rare negative US more negative (i.e., in the positive ecology; M = -0.17, SD = 1.10) compared to CSs paired with frequent negative USs (i.e., in the negative ecology; M = -0.01, SD = 1.16). This joint increase results in the overall main effect that participants rated CSs less positive in the positive ecology (M = 0.09, SD =1.15) than CSs in the negative ecology (M = 0.36, SD = 1.20).

US valence memory. We coded correct answers with “1” and incorrect answers with “0” (including “don’t know” answers) and computed a Welch two-sample t-test. There was no difference between valence memory for frequently and infrequently paired CSs, $t(363.1) = -1.57$, $p = .116$. Participants correctly assigned 85% of frequent CSs and 90% of rare CSs.

US identity memory. Again, the Welch two-sample t-test showed no differences between frequently and infrequently paired CSs, $t(277.73) = -1.35$, $p = .178$. Participants correctly assigned 98% of frequent CSs and 99% of rare CSs.

6.7.3 Discussion

Experiment 3 replicated the results of Experiment 2 while controlling for the rating confound. Even with sampling one CS from the four CSs, we again found the predicted main effect for ecology. Participants rated positively paired CSs more extreme in the negative ecology. Conversely, they rated the negatively paired CS more extreme in the positive ecology.

However, Experiment 3 failed to provide evidence on the memory front. The valence and identity memory tasks were apparently too easy, given the high percentage of correct answers. However, the post-hoc prompted recognition and valence memory tasks do not conclusively allow ruling out memory as a factor during CS ratings (see Gast, 2018).

6.8 General Discussion

Previous research found that negative information outweighs positive information (cf. Baumeister et al., 2001; Rozin & Royzman, 2001): Bad is stronger than good. In our lab, we observed that this point seems not to hold for a paradigm that should show such a negativity bias, namely EC. Negative USs should have a stronger impact on learning. A re-analysis of the meta-analytic data by Hofmann et al. (2010) also showed no evidence for a negativity bias.

To address this lack of a negativity bias in EC, we presented two distal explanations. An evolutionary explanation assumes advantages of negative information because it is adaptive to prioritize negative information, while cognitive-ecological explanations assume basic cognitive processes that interact with the structural properties of the information ecology. In short, negative information may have an advantage because it is less frequent in standard ecologies compared to positive information (see Unkelbach et al., 2019). Based on the latter explanation, we hypothesized that there is no negativity bias in EC because of the equal frequency of positive and negative information in virtually all EC experiments.

Thus, we manipulated the frequency of CS-US presentations in four experiments to test whether we could create a negativity bias based on the differential frequency of positive and negative information. Of note, the cognitive-ecological explanation predicts symmetrical biases. In other words, in a world full of poisonous berries, the rare digestible berries should enjoy a positivity advantage.

In Experiments 1a and 1b, we investigated our assumption by manipulating the frequency of wins and losses in a lottery. Experiments 2 and 3 also used positive and negative pictures as USs. Across all four pre-registered experiments, we found evidence for the predicted evaluation biases.

6.8.1 Relating the Data to the Explanations

As stated above, we presented two distal explanations for negativity biases that differ in how they explain such biases. However, evolutionary and cognitive-ecological explanations might not differ on a proximal cognitive level. For example, both predict attentional advantages (e.g., Pratto & John, 1991; Unkelbach et al., 2019) and deeper encoding (e.g., Baumeister & Cairns, 1992; Unkelbach et al., 2020).

The present results nevertheless show a distinct effect that is difficult to explain from a purely evolutionary perspective, namely the stronger influence of positive USs when positive outcomes or pictures are rare.

However, these data do not imply that evolutionary explanations are wrong. On the contrary, given the substantial evidence for negativity biases (e.g., disgust), dismissing this explanation would be unwise. As often in psychology, several explanations may contribute to substantial and broad phenomena. In the case of EC, though, the nature of the task might prevent the prioritization of negative information, while the cognitive-ecological explanation is ineffective because positive and negative USs are usually equally frequent. If one adapts the frequency distribution of a standard ecology, then one may observe biases predicted by the cognitive-ecological explanation.

A final consideration is why there are negativity biases in paradigms that also use equal frequencies of positive and negative information. For example, Pratto and John (1991) showed that undesirable personality traits (e.g., mean, sadistic) attract more attention than desirable personality traits (e.g., honest, kind), even though they controlled for valence, arousal, word length, and word frequency, and the frequency of occurrence in their experiments was constant across valence.

However, such experiments differ from EC experiments on another level, namely in the stimuli they use. For example, Pratto and John (1991), Anderson and Lampel (1965), and Fiske (1980) used traits or social behaviors. People ontogenetically learn the distributions of such traits: "A basic form of learning available to infants is sensitivity to the frequencies with which events occur in the environment" (Lany & Saffran, p. 232). Consequently, these experiments mirror features, situations, and behavior patterns people encounter daily. In EC, however, the USs consist, for example, of visual or auditory stimuli (Hofmann et al., 2010) that have no such underlying learning distributions. It seems unlikely that people learned the underlying distribution of IAPS pictures or Pixaby sound effects.

A direct test of this assumption follows if one assumes impression formation as a form of EC, an argument made by several authors (e.g., Moran et al., 2016; Navon & Bar-Anan, 2023). Indeed, if traits are used as USs (i.e., by 'pairing' a person with a trait), typical negativity biases appear, as negative information dominates positive information in impression formation tasks (e.g., Anderson & Lampel, 1965; Ronis & Lipinski, 1983).

6.8.2 Limitations and Open Questions

The presented experiments provide first evidence for evaluative biases in EC based on a cognitive-ecological explanation. However, there are some limitations and open questions we plan to address in future research.

First, we only scratched the surface of potential proximal cognitive mechanisms. Various mechanisms might explain why differences in the distribution of stimuli lead to the observed biases. We only used a memory test to investigate whether valence and identity memory might explain more extreme ratings for rare CSs, which failed due to ceiling effects in participants' performance. However, several other mechanisms may lead to the observed biases, such as contrast effects, figure-ground effects, or habituation to the frequent US valence.

Second, we only used words and pictures as USs. However, in EC research, many other USs are used (e.g., flavors, Baeyens et al., 1995; sounds, Moran & Bar-Anan, 2013; smells, Ruszpel & Gast, 2020). Additionally, we did not test our assumptions outside of the laboratory and online environment. Thus, while we tested an explanation and not an effect, future research may investigate the observed biases in different settings and with different stimuli.

Third, we did not test the evolutionary and the cognitive-ecological explanation against each other but only tested the latter. For example, one may test the former by using USs with more biological relevance than the positive and negative pictures frequently employed in EC research.

Finally, we must concede that the observed biases require substantial power to detect them. Across our four experiments, we obtained $F(1, 68) = 6.68$, $F(1, 133) = 8.99$, $F(1, 188) = 19.56$, and $F(1, 187) = 6.37$. As stated, a *p-curve* analysis of these F-values places our average power at 79%, which aligns with our two failures to obtain the effect (i.e., at least 1 out of 5 should probabilistically show no effect). The good news is that the p-curve also shows that our data has evidential value, and the evidential value is not inadequate (see Simonsohn et al., 2015).

6.9 Conclusion

The assumption that bad is stronger than good does not hold true for EC. We propose that the lack of negativity bias follows from the equal frequency of positive and negative information in most EC designs. Thus, typical EC designs do not mirror the ecological information structure that elicits a negativity bias in the first place. Manipulating the frequency of positive and negative information led to the expected negativity bias. In addition, as predicted by a cognitive-ecological approach, it also led to a positivity bias when positive information was rare. Thus, our data offers a partial explanation for the absence of a negativity bias in EC experiments and supports a cognitive-ecological explanation of negativity biases in addition to well-established evolutionary explanations.

Statement of Limitations

Internal validity. All data stems from highly controlled experiments; thus, we believe there are few limitations to internal validity. Experiment 3 explicitly addresses one potential confound.

Construct validity. We only use constructs with standard measures that are frequently employed in this research paradigm. The main variable of frequency is directly measurable. Construct validity is thereby high.

External validity. The setup is artificial in all experiments to test an explanation that predicts negativity biases in learning. It is an open question whether these results replicate in more naturalistic settings, and external validity is thereby low.

Statistical validity. We pre-registered hypotheses and sample sizes for all experiments. We only used standard analytic techniques that are easily reproducible. A *p-curve* analysis (see Simonsohn et al., 2015) of the central main effect shows evidential value. We believe statistical validity to be high.

Statement of author contributions

The authors were equally involved in conceptualizing and designing the experiments. The first author provided the data analyses. The first author wrote the first draft of the manuscript. The final version of the paper is a joint work of both authors. Data was collected by research assistants at a university campus (Experiments 1a, 1b, and 3) or online via Prolific (Experiment 2). The first author was responsible for the Prolific data collection.

Chapter 7: General Discussion

In the previous three chapters, I investigated the impact of organism and environment on EC. Whereas the relevance of environmental features for EC is well-documented (e.g., Hofmann et al., 2010; Hütter et al., 2012; Jones et al., 2010), this focus on the environment followed from the behavioristic tradition of learning psychology. However, this perspective overlooked the organism's influence on the perception of co-occurrences as pairings. This is not to say the organism was *per se* neglected in EC research, it was primarily linked to the *how* question of EC (i.e., the question about EC's underlying processes). However, as the organism has many degrees of freedom to construe the social world (e.g., Kunda, 1990), its impact on stimulus co-occurrences should not go unnoticed. By linking environment and organism, it should be possible (a) to answer the *when* question of EC (i.e., when do people learn evaluations from co-occurrences of two stimuli?), and (b) to explain the lack of negativity bias in EC. I addressed both issues in the previous chapters of this thesis.

Regarding the first point, we proposed that co-occurrences are translated into pairings based on Gestalt Psychology's law of proximity: Two stimuli that appear close together are mentally grouped together (Wertheimer, 1938). Most EC experiments follow this law and present US and CS in close proximity. This spatial proximity is typically supported by the temporal configuration of EC trials (i.e., CS-US, next trial), which mirrors the auditory organization described by Wertheimer: "tap-tap, pause, tap-tap, pause" (p. 74). As following such a rhythmic pattern is almost inevitable (Neisser, 2014), a typical EC experiment's spatial and temporal configuration strongly implies the encoding of CS and US into the same mental episode. Even though the design of the EC experiments presented in Chapter 4 follows a somewhat different structure (i.e., tap-pause-tap, longer pause), the temporal and spatial distance was kept constant between conditions. Thus, it was the judgment that disrupted the encoding of stimuli into the same mental episode, not the change in the structural setup. Functionally, the mental episode resembles an *event file* (Hommel, 2004) or an *event model* (Radvansky & Zacks, 2017). Based on Fiedler's dual-force model (2001; Fiedler & Bless, 2000), we introduced a dual-force perspective of EC, that includes both, stimulus-driven bottom-up input and organism-driven top-down input as influencing factors of mental episodes (cf. Chapter 4). The dual-force perspective assumes that for CS and US to become a pairing, environment and organism must interact. More specifically, we predicted that a top-down influence may prevent or promote the construal of spatio-temporal co-occurrences as pairings. This prediction is not trivial, as spatio-temporal contiguity is probably the strongest predictor for the encoding of CS and US into the same mental episode. Such grouping of stimuli is assumed to be fast, attention-independent, low-level, and bottom-up (e.g., Beck, 1975; Beck & Palmer, 2002; Treisman, 1986). Thus, the interesting test is created by a situation in which the spatio-temporal contiguity of CS and US does not lead to such an encoding due to top-down influences. As was delineated in Chapter 4, standard EC models would not predict an EC effect in this experimental situation. To create the experimental situation, we used Fiedler and colleagues' (2005) setup. Fiedler and colleagues (2005) showed that word recognition might be facilitated by presenting associated pictures (e.g., a man stabbing a woman facilitates recognition of the word 'aggressive'). However, if participants had to judge the picture first (i.e., by answering the question 'Is this behavior aggressive?'), the priming effect was considerably reduced. The authors assumed that judging the behavior closes the mental episode, whereas priming effects require an open mental episode into which the prime (i.e., the picture) and target (i.e., the word) can be encoded. Put differently, the top-down judgment hampered the bottom-up presentation as well as the judgment facilitation for the stimuli.

We found similar results when translating this setup to EC experiments (cf. Chapter 4). Judgments in-between stimulus presentation prevented the encoding of CS and US into the same mental episode. Informing participants that CS and US represent a pairing, however, led to encoding the stimuli into the same mental episode, despite an in-between question. This finding was independent of question type (i.e., location vs. valence) and conditioning procedure (i.e., forward vs. backward). We provided further evidence for the importance of top-down input in Chapter 5 by demonstrating that the conceptual manipulation of CS and US (fit vs. non-fit) influences the strength of EC effects; at least, if the experiment takes place in the laboratory (vs. online). For CS-US combinations that conceptually fit (e.g., rabbit – eggs; conceptual fit: easter bunny – easter eggs), EC was stronger than for combinations that did not fit (e.g., rafting people – eggs). Yet, manipulating the bottom-up input did not yield the same encouraging results as did the top-down manipulations; manipulating the background information of CS-US co-occurrences did not lead to differences in the strength of the EC effect (cf. Chapter 5). Neither brackets nor colors showed to have a significant impact. However, an exploratory analysis showed a slightly stronger EC effect if bottom- up input supported (vs. did not support) the co-occurrences.

In addition to the *when* question, we raised the question of why there is no negativity bias in EC (Chapter 6). From an evolutionary perspective (e.g., Baumeister et al., 2001), it seems to be rather surprising that the re-analysis of the meta-analytic data presented in Chapter 6 did not show a negativity bias in EC experiments, as their structural properties meet all prerequisites (i.e., equal amount of positive and negative information). However, following the EvIE model's rationale, it should become evident why the experimental structure of EC experiments does not lead to valence asymmetries. The EvIE model (Unkelbach et al., 2019) assumes that the organism's evaluation depends on the environment's structure (i.e., the substance). In the real world, this substance consists of many positive and few negative information. The *rare negative information* then enjoys processing advantages (Unkelbach et al., 2020), eliciting a negativity bias. However, the EvIE model also predicts a positivity bias for *rare positive information*. If the substance consists of frequent negative and infrequent positive information, this difference is mirrored in the organism's subjective experience, thus conceding the same processing advantages otherwise attributed to negative information to positive information. Based on these assumptions, we manipulated the ecology of EC experiments. In ecologies with prevalent positive information, EC effects were stronger for negative information. Conversely, in ecologies with frequent negative information and rare positive information, EC effects were pronounced for positive information. Thus, the findings of Chapter 6 aligned with the EvIE model's prediction regarding frequency.

The *when* question and the question of why there is no negativity bias in EC should not be considered separately from each other. On the contrary, both questions are, at least in a broader context, are interconnected. As Fiedler and Unkelbach (2011) noted:

> While preparedness exemplifies a top-down effect, it refers almost exclusively to a priori rules for the learning of fear and phobias (Öhman [& Mineka], 2001), and supposedly reflects stable biological differences between species and individual organisms. It hardly applies to short-term influences of encoding schemes or constructions of CS-US pairs. (p. 640)

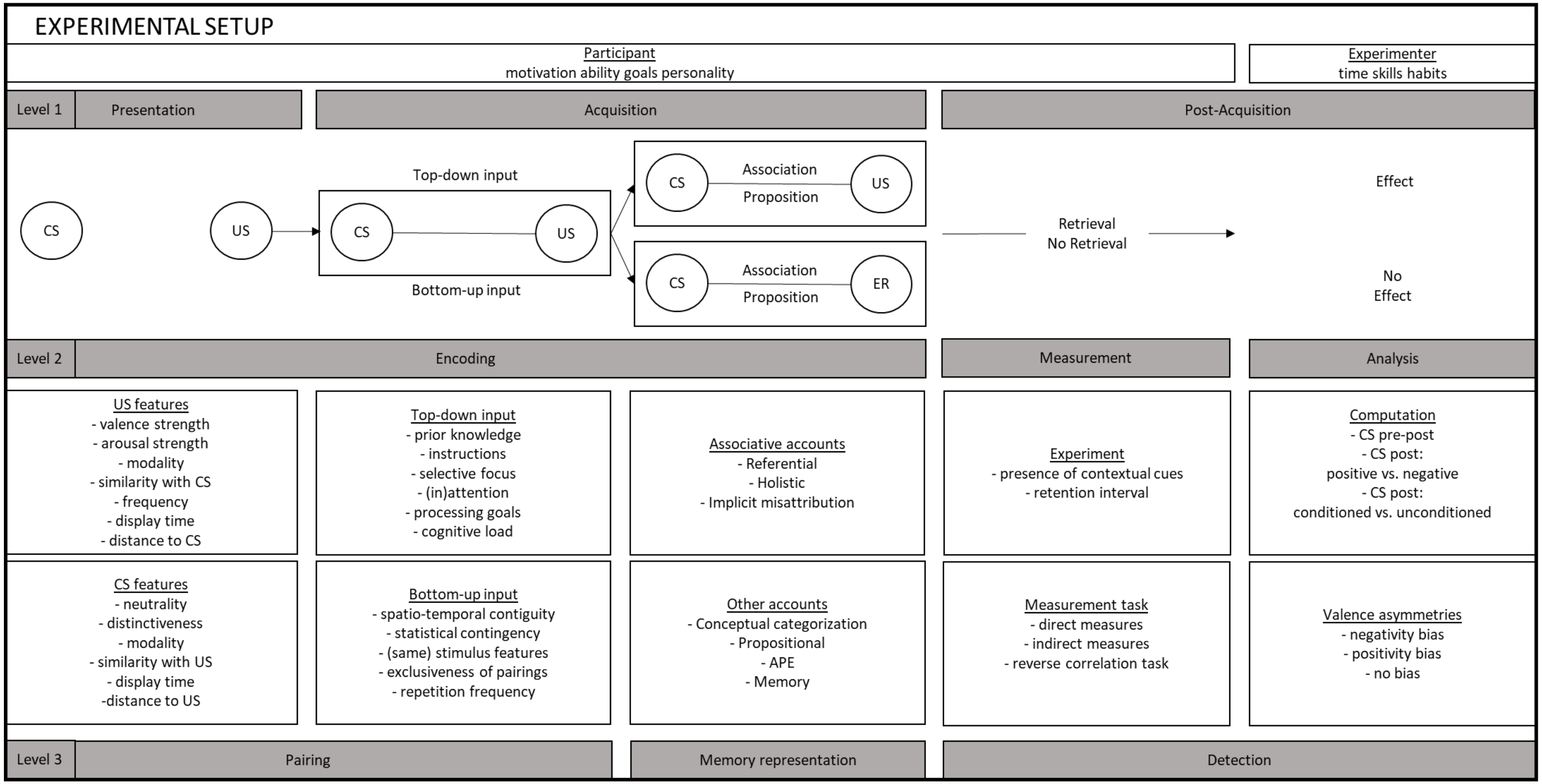
Applying the differentiation between top-down and bottom-up to the question of why there is no negativity bias in EC, preparedness for negative stimuli might be regarded as top-down influence while the frequency manipulation depicts a bottom-up influence. Thus, in a broader sense, preparedness might be expressed as existing knowledge structures (i.e., top-down input), whereas differences in frequency of stimuli account for bottom-up input. Yet, not all top-down processes are determined by the phylogenetic principle of preparedness (cf. Fiedler & Unkelbach, 2011). Additionally, not every variation in bottom-up provides meaningful information regarding the *when* question. For example, even though manipulating the frequency of stimuli represents some form of bottom-up input manipulation, such a manipulation does not answer the question of *when* CS and US are encoded into the same mental episode; yet, it has an impact on the EC effect. Thus, in a next step, I will integrate the findings of this thesis as well as previous research findings into a model of the experimental setup of EC studies.

7.1 An Integral Model of the Experimental Setup of EC Studies

Based on the research introduced in the previous sections, I put forward an integral model of EC experiments (see Figure 7.1). The model illustrates the experimental setup of EC experiments and consists of three different phases on three different levels.

Figure 7.1

An Integral Model of the Experimental Setup of EC Experiments.



Note. The model illustrates three levels that consist of three phases (grey rectangles). All levels and phases consist of the same structural setup; however, dividing this setup into different fragments might facilitate communication about different proceedings during the experimental procedure. The circles depict the physical or conceptual presence of CS and US. The white rectangles around the circles represent the mental episodes in which CS and US are encoded during pairing. The white rectangles on the bottom describe the content of the phases in more detail.

Every level consists of different phases, that are influenced by the participant or the experimenter. The participant's influence in form of motivation (e.g., Veltkamp et al., 2011), ability (e.g., Hasford et al., 2018), goals (Verwijmeren et al., 2012), or personality (e.g., Ingendahl & Vogel, 2023) is thought to eventually influence every phase of the experimental setup (except for the analysis phase). The experimenter's influence is discussed further below. On the most basal level, an EC experiment is composed of the three phases presentation, acquisition, and post-acquisition[33]. First, CS and US are presented, then different processes take place that might or might not encode CS and US into the same mental episode. If encoded into the same episode, links between CS and US are formed, that may require retrieval (e.g. memory traces) or not (e.g., associations). Finally, based on these previous steps, an EC effect might or might not occur. However, EC can also be viewed from a different level. This second level consists of encoding, measurement, and analysis. The *encoding* phase includes both, presentation and acquisition of CS and US. In addition to the features of CS and US, bottom-up and top-down input are encoded. Successful encoding of CS and US does not only force the two stimuli into the same mental episode but also requires the formation of some kind of connection between CS and US or CS and response. In the *measurement* phase, the strength of these (possibly retrieved) connections is measured, which will then be analyzed (i.e., the *analysis* phase). On a third level, the *pairing* of CS and US is dependent on the stimuli's presentation and the first part of acquisition (i.e., the encoding of CS and US into the same mental episode). The *memory representation* phase is then concerned with the formation of associative and/or propositional links, or memory traces between CS and US and/or CS and response. In the *detection* phase, the change in the liking of the CS is measured. This change in liking can also depend on factors like the presence of contextual cues and the retention interval following the acquisition. Additionally, the detection phase contains the computation of the EC effect as well as different outcomes of this computation (e.g., valence asymmetries). Distinguishing between different levels might help with communication about

[33] Different from the basic structure presented in Chapter 2, I do not integrate a special pre-acquisition phase. However, the stimulus features in the presentation phase partly depend on the selection of CS and US (e.g., valence strength/neutrality).

EC experiments. For example, the *when* question is part of the pairing phase, whereas the *how* question is represented by the memory representation phase; yet, that both questions are not fully independent becomes evident when considering the encoding phase on the second level. In the following, I will discuss EC literature in terms of the model's third level. This level consists of (a) pairing, (b) memory representation, and (c) detection, and best illustrates the studies presented in this thesis.

7.1.1 Pairing

As noted before, the co-occurrence of two stimuli is not to be confused with the pairing of such stimuli. In the first step of an EC experiment, CS and US co-occur (i.e., they are presented together, but their distance can vary; cf. Gast et al., 2016 for temporal contiguity; Jones et al., 2009, Experiment 3 for spatial contiguity). In a second step, CS and US must be encoded into the same mental episode to be considered as a pairing[34]. The model thus presents the pairing procedure subdivided into two steps: the co-occurrence of CS and US (left panel of the pairing phase on 'Level 3') and the integration of CS and US features into the same mental episode (right panel; see also Walther et al., 2018). I will first specify some stimulus features that could influence the EC effect before briefly elucidating the potential influence of bottom-up and top-down input for EC.

Stimulus Features. In the first step, the focus is on features of CS and US (for a review, see Hofmann et al., 2010; Moran et al., 2023). Features of the US include the strength of valence (e.g., Gast & Rothermund, 2011a) and related properties (e.g. arousal; Gawronski & Mitchell, 2014), modality (auditory, e.g, Bar-Anan & Dahan, 2013; electrocutaneous stimulation, e.g. Vansteenwegen et al., 2006; haptic, e.g., Hammerl & Grabitz, 2000; olfactory, e.g., Todrank et al., 1995; flavor, e.g. Wardle et al., 2007; verbal sensical, e.g., Zanon et al., 2014; visual, e.g., Luck & Lipp, 2020), perceptual similarity with the CS (e.g., Martin & Levey, 1978), frequency (cf. Chapter 6), display time (i.e., sub- vs. supraliminal; e.g., Heycke et al., 2017; Heycke & Stahl, 2020; Stahl et al., 2016), and distance to the CS (e.g., Jones et al., 2009). Features of the CS, on the other hand, include neutrality

[34] The pairing phase is not restricted to a single pairing procedure; instead, it might also include a counterconditioning (e.g., Gast & De Houwer, 2013) or invalidation procedure (e.g., Sherman & Kim, 2002).

("CSs in EC research are usually neutral in valence", Moran et al., 2023, p. 261), distinctiveness[35] (i.e., "how novel it is for the participant and how unexpected in a given context", Gast, 2018, p. 10), modality (auditory, e.g, Bolders et al., 2012; haptic, e.g., Fulcher & Hammerl, 2001; olfactory, e.g., Rozin et al., 1998; flavor, e.g. Zellner et al., 1983; verbal nonsensical, e.g., Stahl et al., 2009; verbal sensical, e.g., De Houwer et al., 1994; visual, e.g., Gast & Rothermund, 2011a), perceptual similarity with the US (e.g., Davey, 1994), display time (i.e., CS-only presentations; e.g., Moran et al., 2020), and distance to US (e.g., Gast et al., 2016). All these features might or might not be perceived and used by the organism during the second step (i.e., the right panel of the pairing phase on 'Level 3').

Dual-Force Perspective. Thus, whether and when these features are used is part of the second step of the pairing phase. In this step, bottom-up and top-down input influence whether co-occurring stimuli become a pair or not. For example, attending the stimuli's co-occurrence, emphasizing their distinctiveness, or enriching the bottom-up input with prior knowledge should decrease the need for bottom-up input. At the very end of the spectrum, it is possible to conceive of a scenario where no bottom-up regularities are needed to encode CS and US into the same mental episode. For example, merely thinking about a physically absent object might increase confidence in the truth of one's own hypothesis (Koehler, 1991), and intensify the attitude towards the object (Tesser, 1978). Previous research already demonstrated that people can construe two stimuli as a pair even if they do not appear in spatial (or temporal) proximity (e.g., Gast & De Houwer, 2012; see also Moran, Van Dessel et al., 2021). Situational features thereby may make the construal of a pairing more likely, but theoretically, people should be able to construe any two stimuli as a pair, although given the environment's structure, they rarely do so spontaneously. However, most of the time, both bottom-up and top-down influences will interactively determine the construal of the stimuli into a pairing. This assumption differs from previous definitions of pairings, that considered two stimuli paired simply due to their appearance in close spatio-temporal proximity. Additionally, because

[35] Gast (2018) also suggests that concreteness influences memory and thus EC. However, as this influence is not described in more detail (compared to distinctiveness), and because a related concept (i.e., verbal sensical) is depicted in the modality feature, I decided not to list concreteness as a distinct factor.

instructions typically do not provide any indication of the meaning behind this 'pairing', participants were regarded as passive observers of this presentation (Hütter, 2022). Based on the dual-force perspective, however, stimuli are considered paired when they are encoded into the same mental episode; a process that accredits a more active role to the participant (cf. Chapter 4).

7.1.2 Memory representation

In Chapter 2, I distinguished between associative and non-associative accounts based on their assumptions about the links that are formed during conditioning (i.e., memory trace, association and/or proposition). If links are formed between CS and US, they represent an instance of S-S learning. If they form between CS and response, they depict S-R learning. As the response in EC is evaluative in nature, I labelled it *evaluative response* (ER) in Figure 7.1. Because EC effects were found for both, S-S learning and S-ER learning (e.g., Gast & Rothermund, 2011b), some researchers argue that EC might depend on both, S-S and S-ER learning (Sweldens, 2018).

Distinguishing between S-S and S-ER learning as well as between the different accounts introduced in Chapter 2 (see also Figure 7.1) allows different predictions about some of the stimulus features mentioned earlier. For example, the implicit misattribution account suggests that mildly valenced USs have an advantage over strongly valenced stimuli (Hofmann et al., 2010). Conversely, the US-Valence Hypothesis (S-S learning) would predict that strongly valenced USs induce greater EC effects (Gast & Rothermund, 2011a). Additionally, some accounts assume that EC effects depend on perceptual similarity of CS and US (e.g., implicit misattribution, Jones et al., 2009; conceptual categorization, e.g., Davey, 1994). Likewise, the implicit misattribution account suggests that spatial proximity influences EC. Additionally, most non-associative accounts assume that EC is dependent on the number of CS-US co-occurrences (i.e., repetition frequency).

7.1.3 Detection

Like the pairing phase, the *detection* phase consists of two steps. The EC effect can be influenced by the measurement task as well as experimental factors like the presence of contextual cues during measurement and retention interval. Additionally, the EC effect itself can be computed in a variety of ways. Detection of valence asymmetries is dependent on the type of computation.

Experiment. Whether an EC effect is detected or not as well as the strength of this effect might depend on experimental circumstances. For example, context cues (i.e., cues that are present during learning and measurement) might facilitate retrieval and should thus lead to stronger EC effects (Gast, 2018).

Additionally, although typical EC experiments measure the EC effect directly after the acquisition phase, some researcher implemented a retention interval (e.g., Gast et al., 2012; Förderer & Unkelbach, 2013; Fulcher & Cocks, 1997)**.** The retention interval is defined as the time span between learning and measurement phase; however, as it influences the measurement, it is allocated to the measurement phase in Figure 7.1.

Measurement Tasks. As EC is per definition the change in the valence of a CS due to its pairing with negative or positive stimuli (De Houwer, 2007), the main dependent variable in EC experiments is the liking of the CS. The liking of the CS can be measured using direct (i.e., self-report) or indirect measures (e.g., IAT). I use the terms *direct* and *indirect* rather than *explicit* and *implicit* based on the recommendations of Corneille & Hütter (2020; see also Greenwald & Banaji's, 2017; for a slightly different perspective see Gawronski et al., 2020). There are different types of indirect measurement tasks (for an extensive overview see Greenwald & Lai, 2020, Table 1). I will explain in more detail one variant of each subgroup: Implicit Association Test (IAT; Implicit Association Test variations), Name-Letter Effect (NLE; other methods), and Evaluative Priming Task (EPT; priming variations). The IAT (Greenwald et al., 1998) is one of the first (Gawronski et al., 2020) and most-cited (Greenwald & Lai, 2020) indirect measures. If used to measure an EC effect, participants in a first phase observe USs (e.g., 'good' and 'bad') together with CSs (e.g., a pictures of neutral faces) presented in the center of a screen. In a second phase, USs and CSs appear on the upper-left or upper-right part of the screen, while positive (e.g., fantastic) or negative attributes (e.g., horrible) must be classified in accordance with US and/or CS depicted in the corner of the screen. In congruent trials, the positively (negatively) paired CS is presented together with the 'good' ('bad') US; in incongruent trials, CS and US do not match (e.g., positively paired CS and 'bad' US vs. negatively paired CS and 'good' US). The block order of congruent vs. incongruent trials is usually counterbalanced (e.g., Schnabel et al., 2008). The EC effect is then represented in participants' response speed in congruent versus incongruent trials (Hütter, 2022). A second measure, often used to measure implicit self-esteem (Stieger & Burger, 2013) but rarely applied in EC research (cf. Hofmann et al., 2010), is the Name Letter Task (Kitayama & Karasawa, 1997; also called Initial Preference Task (IPT), Stieger et al., 2012). This task is based on the Name Letter Effect (NLE, Nuttin, 1985, 1987). The NLE states that people prefer letters of their own names over other letters in the alphabet; this effect is even more pronounced for the initials of one's own first and last name (compared to the remaining letters; Dijksterhuis, 2004). When applied to EC, the IPT was used to measure differences in self-esteem after a conditioning procedure (e.g., presenting the word 'I' together with positive trait words; Dijksterhuis, 2004). Another, more frequently used measure from the family of priming paradigms is the Evaluative Priming Task (EPT; Fazio et al., 1995). In this task, participants are primed with previously conditioned stimuli (that they are usually instructed to ignore). As conditioned stimuli should facilitate the responses on congruent (vs. incongruent trials), higher speed and less errors indicate the EC effect (Hütter, 2022). Another interesting and rather new measure (at least for EC) is the reverse correlation task (Rougier & De Houwer, 2023; also referred to as 'classification image technique', e.g., Brinkmann et al., 2017). The task originated from signal detection theory (Brinkman et al., 2017) and was used before in other areas, for example in approach-avoidance training (Rougier et al., 2021). A basic reverse correlation task is composed of four steps: (a) presenting participants a large number of stimulus material with random variations (e.g., faces superimposed by random noise), (b) asking participants to select the images that match their mental representation, (c) computing so-called 'classification images' based on the participants' responses, and (d) let other participants (so-called 'independent judges') rate these 'classification images' on various features (e.g., personality traits like 'trustworthiness' or 'dominance'; e.g., Dotsch & Todorov, 2012). Such 'classification images' are thought to represent stimulus features that influence social judgment. Therefore, they are considered an approximation of the mental representations that were brought into play (Brinkmann et al., 2017). In the context of EC, participants first observed CS-US presentations. CSs showed blurred pictures of two neutral faces labelled 'Andy' or 'John', and USs were negative and positive pictures taken from IAPS. After this acquisition phase, participants were instructed to select the face that is most similar to the face of John or Andy. Based on these responses, 'classification images' were computed. In a last step 'independent judges' evaluated the 'classification images' on various personality traits, liking and positivity (Rougier & De Houwer, 2023). Thus, in contrast to other measurement methods used in EC experiments, the reverse correlation task is not restricted to evaluations of valence, but also allows the evaluation of social dimensions.

Computation and Valence Asymmetries. For the computation of EC effects, it is useful to differentiate between the different kinds of CSs: positively paired CSs (CS_{pos}), negatively paired CSs (CS_{neg}), and non-paired CSs (CS_{non}). There are many different ways of computing the EC effect (see Moran et al., 2023, Figure 2, for a graphical illustration of EC effect computations). A distinction can be made between three main procedures: CS_{pre} and $CS_{pos/neg}$, $CS_{pos/neg}$ and CS_{non}, or CS_{pos} and CS_{neg}. In experiments with baseline ratings of CSs (CS_{pre}), the EC effect can be calculated by subtracting post-ratings of CS_{neg} from CS_{pre} or by substracting CS_{pre} from post-ratings of CS_{pos} (within participants). In a similar vein, the EC effect can be calculated by substracting post-ratings of CS_{neg} from ratings of CS_{non} or by substracting ratings of CS_{non} from post-ratings of CS_{pos} (within/ between participants). Additionally, the EC effect can be calculated using negative and positive paired CSs only, subtracting CS_{neg} from CS_{pos} (within/between participants). Yet, this latter computation of EC effect does not allow to detect valence asymmetries. Thus, even though the EC effect provides information about the change in CS liking, it does not represent differences in the liking for CS_{pos} compared to CS_{neg}.

7.1.4 Implications of the Integral Model

Even though the EC setup can be subdivided into different phases, the phases are thought to interact and, in a final step, elicit an EC effect (or not). Thus, in the next section I review EC literature in terms of the integral model introduced earlier with a focus on the impact of these findings on the EC effect. Particular attention is paid to the reinterpretation of findings in relation to the dual-force perspective.

Stimulus Features and EC effects. First, features of CS and US might influence the strength of the EC effect. For example, previous research found that normed ratings of US valence (e.g., of IAPS pictures) might not always correspond to participant's ratings of US valence (e.g., Ingendahl, Woitzel, Propheter et al., 2023). For example, individuals with higher scores on 'Agreeableness' and 'Neuroticism' exhibited stronger evaluations of US pictures, potentially elucidating the more pronounced EC effects observed in these participants (Ingendahl & Vogel, 2022). Regarding the strength of valence, some researchers provides evidence for strong EC effects with strongly valenced USs (Baeyens et al., 1988), or at least assume that strongly valenced stimuli might have an advantage (Gast & Rothermund, 2011a). However, other research suggests that mildly valenced USs are beneficial (Jones et al., 2009). Although evidence tends towards indicating higher effectiveness of mildly valenced stimuli (De Houwer, 2011), more research is needed to provide a more definitive conclusion. For arousal, research could show that EC effects were stronger for high than for low arousal USs (e.g., Gawronski & Mitchell, 2014; Zerhouni & Lepage, 2018). Additionally, it was demonstrated that arousal can be conditioned in a similar vein as valence (Balas & Pochwatko, 2021). Glaser et al. (2018) also go beyond the scope of pure positivity or negativity by demonstrating the conditioning of ambivalence to CSs. In addition, it has been shown that attributes can be subject of conditioning (e.g., Förderer & Unkelbach, 2011); an effect referred to as *attribute conditioning* (Unkelbach & Högden, 2019). With regard to CS features, distinctiveness is thought to serve as a retrieval cue, thus increasing the probability of EC to occur (Gast, 2018). Similarly, CS salience was shown to increase the EC effect (Jones et al., 2009). Additionally, using invariable (vs. variable) CSs leads to stronger EC effects (Reichmann et al., 2023), whereas EC effects vanish when changing a CS's appearance (Unkelbach et al., 2012). Concerning stimulus modality, it was found that haptic USs are less suitable than visual, auditory, taste/flavor, odor or verbal sensical stimuli (Hofmann et al., 2010); additionally, it might be beneficial to choose CS and US from different modalities (e.g., Blask et al., 2012; Gast et al., 2016; but see also Hofmann et al., 2010). Similarly, Alves and Imhoff (2023) could show that stimuli from different categories (e.g., animals used as USs and Kanji used as CSs) lead to assimilative effects (i.e., standard EC effects), whereas stimuli from the same category (e.g., faces used as USs and CSs) were shown to elicit contrastive effects. Additionally, spaced presentations of CS-US pairing might have an advantage compared to massed stimulus presentations (Richter & Gast, 2017). Additionally, stronger EC effects were reported for shorter temporal intervals between CS and US (e.g., Gast et al., 2016) or CS and ER (Blask et al., 2020), and higher spatial proximity of stimuli (Jones et al., 2009; Hughes, Mattavelli et al., 2018; Rydell & Jones, 2009; for an overview see Moran et al., 2023).

The relationship between stimulus features and bottom-up (and top-down) input becomes evident with regard to this spatio-temporal proximity of stimuli. Distance and display time are important factors to distinguish between contingency and contiguity (cf. Chapter 1) as prerequisite for EC. Some researchers have suggested contingency, rather than contiguity, as the driving factor behind EC effects (Hütter et al., 2014). Consistent with the assumption, reduced EC effects have been found for US-alone (e.g., Hammerl et al., 1997) and CS-alone presentations (Halbeisen & Walther, 2016). However, while contiguity and contingency might be confounded in most EC experiments (Hütter et al., 2014), research that has disentangled both factors showed EC to be insensitive to contingency (Baeyens et al., 1993). This latter finding is supported by meta-analytic data (Hofmann et al., 2010). Thus, contingency was discarded as prerequisite for EC, and researchers relied on contiguity instead (e.g., Martin & Levey, 1994). Still, research demonstrating the significance of contiguity focusses primarily on the spatial and temporal parameters of CS and US presentations (e.g., Kattner et al., 2012; 2014). While our findings in Chapter 5 support the importance of spatio-temporal contiguity, Chapters 4 and 5 also highlight the necessity of considering the influence of top-down input on stimulus co-occurrences. Thus, during the second step of the pairing phase, the model does not only include bottom-up input but also integrates top-down input, enabling interaction between both forces.

Dual-Force Perspective and EC effects. In typical EC experiments, CS and US are presented in close spatio-temporal proximity, without giving an instruction about the purpose of their co-occurrence (Hütter, 2022). Thus, the strong bottom-up input together with the weak top-down input leads to the formation of a joint mental episode of CS and US (cf. Figure 4.1, left panel). This crucial role of bottom-up input was already presented by Hughes, Mattavelli and colleagues (2018). The authors found larger EC effects for CS and US presented in closer proximity, as opposed to when they were presented farther apart (see also Jones et al., 2009; Rydell & Jones, 2009). These findings are supported by the results of Chapters 4 and 5. Even though reduced by the top-down input, strong bottom-up input led to an EC effect in Chapter 4. Similarly, adding bottom-up input when spatio-temporal contiguity is given and top-down input is absent seemed not to contribute further to the encoding of CS and US into the same mental episode. Additionally, we presented first evidence for the influence of top-down on EC effect in Chapter 4. Whether participants encoded CS and US into the same mental episode depended on the type of top-down input: supporting top-down input led to encoding, whereas preventing top-down input hampered encoding. Additionally, the findings regarding top-down input presented in Chapter 5 suggest that EC effects increase when CS-US pairings are enriched by existing knowledge. Similarly, previous research showed that EC can also occur when both stimuli are only indirectly related (De Houwer, 2011), for example, if prior knowledge about a CS-US link exists (Hammerl & Grabitz, 1996).

In addition to these findings, the dual-force model might be able to explain failures of finding EC effects using subliminal or incidental EC paradigms. EC is classified as subliminal if a CS is presented less than 50 ms (Hofmann et al., 2010). Additionally, factors like stimulus size, contrast, and masks should be considered in the context of subliminal CS presentation (Stahl & Bading, 2020). In previous research, the absence of EC effects in subliminal paradigms was discussed in terms of contingency awareness for CS-US pairings (e.g., Heycke et al., 2017; Heycke & Stahl, 2020; Stahl et al., 2016). However, if interpreted in terms of the dual-force perspective, subliminal EC (e.g., very short stimulus presentation) might depict a rather weak bottom-up input (see also Högden et al., 2018). Together with weak input from the organism’s side (Högden et al., 2018), chances are that CS and US are not encoded into the same mental episode. Högden et al. (2018), however, in line with Dehaene et al. (2006), suggested that strong bottom-up input combined with strong top-down input leads to consciousness (i.e., awareness). Contingency awareness is usually interpreted as evidence for a propositional link between CS and US. The dual-force perspective, on the other hand, does not rely on underlying processes, but only predicts the encoding of CS and US into the same mental episode when there is strong bottom-up and top-down input.

A very similar argument can be applied to incidental EC (e.g., Olson & Fazio, 2001; see also Moran, Hughes et al., 2021). A task qualifies as incidental if (a) participants are neither asked to learn nor to attend the CS-US combinations, (b) an unrelated task distracts participants on a moderate level, and (c) the presentation of CSs and USs appears to be part of an apparently random sequence of CS-US combinations, blanks, individual stimuli, and pauses (Stahl & Heycke, 2016). This implementation of weak bottom-up and top-down input might thus explain the reduction of EC effects in these experiments. Specifically, even though CS and US are labeled as “pair” (Olson & Fazio, 2001, p. 414), they might not have been encoded into the same mental episode (i.e., they do not qualify as pairing based on the requirements introduced in Chapter 2).

However, there are also experiments that might be considered as providing strong top-down input. For example, Corneille and colleagues (2009) found stronger EC effects for participants with a similarity (vs. difference) focus. Even though the authors interpreted their results in terms of memory representations (i.e., underlying processes of EC), they can also be explained from a dual-force perspective. Encoding of CS and US might be facilitated by directing their attention to similarities of CS and US (supporting top-down input) and hampered by the difference-focus (disrupting top-down input). The fact that the EC effect was weakened (and did not completely diminish) was attributed to the difficulty of activating “a pure difference focus processing mode” (Corneille et al., 2009, p. 281). Thus, it was assumed that participants did process some similarities despite of being in the difference-focus condition. In terms of the dual-force perspective, it is possible that the bottom-up input was stronger than the top-down input for these participants. Similar results were provided by other researchers. Mierop and colleagues (2020) for example showed that top-down input (i.e., cognitive load) reduced EC. Field and Moore (2005) provided evidence that top-down input (i.e., attention) increased EC. The study from Mierop and colleagues (2020) is of special interest, as it shows that EC is reduced under cognitive load even if the individual stimuli (i.e., CS or US) are remembered. Instead, EC depends on the recognition of CS and US as a pairing; thus, this recognition as a pairing (and not the recognition of CS and US per se) is hampered by cognitive load. Put differently, top-down input (i.e., cognitive load) prevents the encoding of CS and US into the same mental episode. Another study by Hughes, Ye and De Houwer (2019) presented participants CS-US pairings together with context stimuli (e.g., identical: day-day vs. opposite: day-night) and found no EC effects for the 'opposite' condition; yet, if the 'opposite' meaning serves as a symbolic cue, one might expect a full reversal of the EC effect (but see also Hughes, Barnes-Holmes et al., 2018). In terms of the dual-force perspective, however, the 'opposite' condition might represent a top-down input that prevents the encoding of CS and US into the same mental episode and thus, an attenuated (instead of a reversed) EC effect would be readily predicted. In more extreme cases, mere instructions about CS-US pairings (without the physical presence of CS and US) were shown to lead to EC effects (Gast & De Houwer, 2012; see also Moran, Van Dessel et al., 2021). In this context, the substantial influence of top-down input might render the bottom-up input obsolete.

Memory Representation and EC effects. Even though the dual-force perspective does not make assumptions about underlying processes of EC, it aligns more cohesively with certain process theories (cf. Chapter 2) compared to others. For example, the implicit misattribution model suggests that EC effects are the result of mistakenly attributing the US evaluation onto the CS (Jones et al., 2009). A critical element of such source confusability is the simultaneous presentation of CS and US (see also Jones et al., 2010; Sweldens et al., 2010). While simultaneous conditioning might lead to EC without contingency memory (Hütter & Sweldens, 2013), EC effects do not differ for simultaneous, forward, and backward conditioning procedures (e.g., Gast et al., 2016; Kim et al., 2016; Mallan et al., 2008; but see also Stahl & Heycke, 2016; Zerhouni et al., 2018). In line with these findings, Chapter 4 shows EC effects using a sequential conditioning procedure. Thus, although the implicit misattribution account might have predicted the results of the experimental condition (no EC effect), it cannot explain the occurrence of EC effects in the control conditions. However, research suggests that sequential conditioning might depend on contingency memory (Hütter & Sweldens, 2013; Stahl & Heycke, 2016). Yet, from a *how* perspective, it does not seem clear why a valence question in-between stimulus presentation should disrupt memory. From a *when* perspective however, the inattention to the subsequently presented CS (or the selective focus on the US) does prevent the encoding into the same mental episode; a prediction that can also be derived from memory models (e.g., Stahl & Aust, 2018). Other accounts might be compatible with our results but are unlikely to underlie EC due to previous findings from other studies. The conceptual categorization account (e.g., Davey, 1994) for example assumes that EC effects depend on perceptual similarity of CS and US; yet, meta-analytic data did not support this assumption (Hofmann et al., 2010). The remaining accounts differ primarily in terms of their predictions regarding contingency awareness (for reviews see e.g., Gast et al., 2012; Moran et al., 2023; Sweldens et al., 2014). Associative accounts suggest that EC can occur without contingency awareness, non-associative accounts do not (cf. Hofmann et al., 2010). The findings of Chapter 4 are more or less compatible with both assumptions. However, the findings of Chapter 4 show that top-down and bottom-up input might be orthogonal to associative and propositional processes (cf. Figure 4.10). From a dual-force perspective, it is rather easy to argue why top-down input reduces the EC effect: Attention might be withdrawn from the CS, preventing the encoding of US and CS into the same mental episode. However, from a mere propositional perspective, it is less clear why a question that highlights the US valence should reduce the EC effect (Experiment 2). Additionally, a propositional account might have difficulties to explain why the timing of the judgment should matter (Experiment 3). On the other hand, assuming that the top-down input is of associative nature, fails to explain why the pairing instruction (Experiment 4) leads to an EC effect despite of the in-between question. However, if CS and US are part of the same mental episode, these episodes might contain associative or propositional links that also have an influence on the EC effect. For example, different relational qualifiers might influence the EC effect to varying degrees (Hughes et al., 2020), or an increase of co-occurrences might increase EC by strengthening associations between CS and US (cf. Hofmann et al., 2010). However, evidence for the latter assumption is inconclusive (cf. Chapter 2). Additionally, memory models do not assume the number of pairings to moderate EC if memory for pairings is perfect (Gast, 2018). Still, disentangling the *when* and the *how* question might help to better understand when EC occurs and when it does not.

Measurement tasks and EC effects. As mentioned before, the dependent variable measured in typical EC experiments is the liking of the CS. Such measurement requires that the CS is, or at least recently was present when participants give their responses. Due to the previous pairing with the US, the CS is assumed to serve as the ideal retrieval cue (Gast, 2018). However, there might be other important retrieval cues (e.g., context cues) that influence the EC effect. Even though not tested in EC paradigms yet, Gawronski and colleagues (2010) showed that the evaluation of a target stimulus depended on the background color of behavioral descriptions presented during an *attitude formation* phase. For example, if the target stimulus was presented with positive descriptions on a yellow background, and the target stimulus was then presented on a yellow background during the measurement phase (using AMP), the target stimulus was evaluated more positive than when presented with a blue background (the background color that was used for negative descriptions of the target stimulus during attitude formation).

Although evidence for the influence of contextual cues is pending, there are some studies investigating the impact of the retention interval on EC effects. In a typical EC experiment, participants liking of the CS is tested right after the acquisition phase (cf. Moran et al., 2023). However, there are some studies that tested EC effects directly after acquisition *and* after retention intervals of different length (e.g., Gast et al., 2012; Förderer & Unkelbach, 2013; Fulcher & Cocks, 1997). These studies found that EC effects were quite stable over time. However, EC effects in the study of Fulcher and Cocks (1997) relied on US identity awareness (cf. Stahl et al., 2009) at the first measurement point (i.e., immediately after acquisition). Conversely, Gast and colleagues' (2012) findings show to be dependent on the presence of contiguity memory at the second measurement point (i.e., remembering CS-US pairings 9-10 days after acquisition). Förderer and Unkelbach (2013), on the other hand, found a decrease in contiguity memory and argue that contingency awareness during the acquisition phase might be necessary for EC to occur. This assumption is further supported by meta-analytical data (Hofmann et al., 2010).

Thus, the question of retention is tightly related to the question of underlying processes of EC. Indirect measures like the IAT are often assumed to measure associations (Corneille & Hütter, 2020). As already mentioned in Chapter 2, propositions are thought to be influenced by assumptions about their truth value, whereas association formation is thought to rely on spatio-temporal contiguity (Gawronski & Bodenhausen, 2006, 2011). However, recent research challenges that assumption. Moran and colleagues (2017) found that only the IAT (but not other priming measures like the EPT) was less sensitive to validity information. Even though EC is larger for self-report than for indirect measures, Hofmann and colleagues (2010) pointed out that EC effects measured with indirect measures differed significantly from zero. Within the indirect measures, the strongest EC effects were assessed with the IAT and the IPT, followed by affective priming paradigms (Hofmann et al., 2010). These findings might be attributed to the assessment of actual associations using indirect measures (Hofmann et al., 2010); however, they do not rule out the possibility of poor psychometric qualities of indirect measures (for a review see Bar-Anan & Nosek, 2014). Indeed, the subject of debate revolves around what indirect measures truly capture[36]. Yet, even though the IAT is criticized as a measure of implicit constructs (Schimmack, 2021; but see also Kurdi et al., 2021), it is currently considered the best method for assessing automatic judgments at the individual level (Vianello & Bar-Anan, 2021). As mentioned earlier, automaticity can be subdivided into awareness, intention, control and efficiency (Bargh, 1994). Additionally, the IAT might also be able to capture non-automatic features (Corneille & Hütter, 2018); however, a detailed discussion of indirect measures would go beyond the scope of this thesis. In EC research, such indirect measures are for example used to "safeguard against influences of demand awareness during measurement" (Förderer & Unkelbach, 2013, p. 382) or "self-presentation" (Richter & Gast, 2017, p. 6).

Yet, findings obtained with direct and indirect measures sometimes dissociate (Stahl & Aust, 2018), and more research is needed to determine the reason behind this dissociation. For example, some studies found stronger EC effects for an increase of stimulus co-occurrences using direct measures (cf. Chapter 2), whereas studies using indirect measures have failed to find such an effect

[36] I will not delve further into this on-going debate about what implicit measures measure (e.g., Brownstein et al., 2019). However, for reviews about implicit measures see for example Gawronski et al. (2020), Gawronski & Hahn (2018), and De Houwer et al. (2009). For recommendations on the future use of implicit measures see van Dessel and colleagues (2020).

(e.g., Hu et al. 2017a, Kurdi & Banaji, 2019). A similar pattern can be observed for extinction, while the reversed is true for blocking effects in EC (see Moran et al., 2023). Additionally, emotion-regulation strategies like suppression, and reappraisal effectively reduced EC on self-report measures, but did not reduce EC when measured with the EPT.

Regarding the reverse correlation task, first evidence was provided by Rougier and De Houwer (2023), who found an EC effect as well as an effect of US valence on socially relevant traits (when controlling for trait valence): in the US positive (vs. the US negative) condition, 'classification images' differed more on traits that were socially relevant (e.g., warmth) than traits that were socially irrelevant (e.g., cleverness, Rougier & De Houwer, 2023).

Computation and EC Effects. In typical EC designs, the mere co-occurrence of CS and US usually results in assimilative effects (i.e., a standard EC effect, Moran et al., 2023). However, there are some exceptions. For example, reversed EC effects can be found when introducing contrastive judgment tasks (e.g., Unkelbach & Fiedler, 2016), relational qualifier (e.g., Förderer & Unkelbach, 2012), or same object categories (e.g., Alves & Imhoff, 2023). Additionally, reversed EC effects were found for unaware pairings (e.g., Stahl et al., 2009). However, this latter finding might be due to the usage of stimuli from pretested databanks (Ingendahl, Woitzel, Propheter et al., 2023; see also Ingendahl & Vogel, 2022). Yet, this case shows that the cleanest way to capture EC might be to present participants with various stimuli in a prerating phase. Based on these idiosyncratic preratings, sufficiently positive and negative stimuli can be selected as USs, while neutrally rated stimuli can serve as CSs. EC effects can then be calculated by computing CS_{pos}-CS_{pre} and CS_{pre}-CS_{neg} (within participants). However, researchers might decide against such a proceeding due to time constraints, a lack of (programming) skills, or out of habit. Additionally, some study designs do not allow for this kind of calculation (e.g., between participant designs; Moran et al., 2023). Yet, meta-analytic data shows that EC effects are larger for idiosyncratically selected USs. The more interesting finding regarding the analysis phase is, however, that EC effects are smaller for CSs with initial valence and CSs assigned on a group level (Hofmann et al., 2010). For this reason, we assessed baseline ratings of CS liking before the conditioning phase in Chapter 5. Although we did not find the predicted result using only postratings (Experiment 1), we found the predicted pattern when adding a prerating and computing the average CS rating by subtracting CS_{pre} from CS_{post} (Experiment 2).

As mentioned before, the calculation type of EC effects also influences the detection of valence asymmetries. Thus, evidence for the lack of negativity bias in Chapter 6 stemmed from calculation of effect sizes based on conditioned and unconditioned CSs (pre-post and within participants). The calculation of the EC effect based on postratings, on the other hand, might prevent the discovery of valence asymmetries. For example, if in a hypothetical condition *positivity bias*, participant's average rating of CS_{pos} is 5, and the average rating of CS_{neg} is -3, the respective EC effect is 5 -(-3) = 8. In a hypothetical condition *negativity bias*, where participants evaluate the CS_{pos} on average with 3 and the CS_{neg} with -5, the EC effect is also 8. Thus, even though there is a negativity and a positivity bias in the respective conditions, this specific calculation of the EC effect is unable to depict this difference. In addition to the magnitude of the EC effect, it might thus be helpful to report average CS ratings (cf. Chapter 6).

7.1.5 Summary

The arguably most cited definition of EC is De Houwer's (2007) functional definition of EC as an effect. EC is thus the change in liking of the CS due to its pairing with a valent US. Until now, the focus has primarily been on understanding the underlying processes of this pairing. The assumption of a mental episode incorporating both CS and US has been taken for granted. *When* CS and US are encoded in such an episode has not been asked in previous research. Similarly, little attention has been given to a more detailed examination of the structure of EC experiments. However, as we demonstrated in Chapters 4-6, both aspects have a significant impact. Based on these findings, an integral model for the structure of EC experiments has been established. The model illustrates several variables that can influence the EC effect. Even though some factors are usually kept constant across EC studies (e.g., the symmetrical presentation of positive and negative USs), others vary immensely (e.g., stimulus material, interstimulus interval, stimulus positioning, Walther et al., 2018). This might be informative regarding EC's robustness and unproblematic for within-comparisons of experiments (as presentation duration is usually kept constant between experiments of the same study). However, it makes a comparison between studies challenging. The model can thus help identify potential differences in EC experiments and support a more systematic approach to future studies. Additionally, it might be helpful to focus on aspects previously considered as given.

7.2 Alternative Explanations

Other researchers also point to the necessity of using more standardized research methods and well-defined experimental designs to facilitate comparison between different EC studies. Walther and colleagues (2018) for example propose “a more theory-driven and systematic variation of encoding-relevant parameters” (p. 4). Their *binding theory* has some parallels with the dual-force perspective (cf. Chapter 4) and is thus examined in more detail in the next section. Another model that has striking similarities with the setup of some experiments in Chapter 5 is the *shared feature principle* (Hughes et al., 2020). The shared feature principle suggests that if two objects share one feature (e.g., color), people will assume that they share other features as well (e.g., valence); the principle is thus described in more detail below. Additionally, even though Chapter 6 provides evidence for a cognitive-ecological explanation, we did not test the ecological and the evolutionary explanation against each other. Thus, I will briefly discuss the role of the evolutionary approach for our findings.

7.2.1 The Binding Theory

A concept very similar to the mental episodes put forward by the dual-force perspective is that of a binding idea (Walther et al., 2018). Based on the feature integration theory (Treisman & Gelade, 1980) and the theory of event coding (TEC, Hommel et al., 2001; Hommel, 2004), the binding perspective assumes that the occurrence of EC depends on the integration of CS and US features into one cognitive representation (Walther et al., 2018). Additionally, the binding approach assumes that the processing of CS and US features is selective and depends on the current goals of the organism. Preliminary evidence that supports this assumption comes from earlier studies performed by Blask and colleagues (2017b). The authors found a significant reduction in EC when participants were asked to (selectively) ignore the US, even though spatio-temporal contiguity was given. In line with the binding perspective, the dual-force perspective assumes that integration of CS and US constitutes a rather early stage of the encoding process. Additionally, like the binding theory, the dual-force perspective considers the organism's input an essential component for encoding CS and US into a mental episode. However, the binding theory suggests that stimulus features are not limited to perceptual characteristics; they also refer to response features (e.g., emotional responses towards the US). Walther and colleagues (2018) compare their binding perspective with the fusion of CS and US in immediate memory as proposed by Martin and Levey (1978), and the implicit misattribution of an affective response to the CS as suggested by Jones and colleagues (2009). However, they assume EC not to be "a failure of the cognitive apparatus", but "an adaptive process tailored toward situational demands" (Walther et al., 2018, p. 7). Thus, unlike the dual-force perspective, the binding theory is guided by the question of *how* EC occurs (even though it does not directly address the structure of EC's underlying processes). Additionally, explaining the results of Chapter 4 requires some additional assumptions. First, one must assume that the experiment's instructions to 'observe the presented pictures attentively' did not translate to the goal of observing both pictures attentively. In this case, the question in-between stimulus presentation might have selectively directed the participants' attention to features of the US (i.e., features relevant for the participant's current goal to answer the question), whereas the irrelevant features of the CS were not encoded. However, one would still have to explain why the timing of the judgment (Experiment 3) made a difference. Thus, the results of Chapter 4 are more parsimoniously explained by the dual-force perspective.

7.2.2 The Shared Feature Principle

The importance of stimulus features is also emphasized in the shared features principle (Hughes et al., 2020). Hughes and colleagues (2020) differentiate between source features and target features, with target features being the "features of an object about which assumptions are being made" and source features giving "rise to assumptions about target features" (p. 2). The respective objects that possess these features are called target object and source object. To exemplify the feature transformation between source and target object, the authors use the halo effect. If a person X (source object) is said to be attractive (source feature), people might assume the person X itself or the person's partner (person Y; target object) to be intelligent (target feature). However, feature transformations also allow for the opposite effect (it might for example be assumed that an attractive woman is not intelligent). For feature transformation to take place, it is required for source and target object to share some feature (i.e., both objects must have a feature in-common that can be identified). To test their assumption, the authors used a setup similar to EC experiments. They showed participants two valenced words (i.e. a positive and a negative US) together with one non-word (CS). In half of the trials, all three words were presented in white on black background. However, in the other trial half, one of the USs was presented in the same color as the CS (e.g., yellow), while the other US was presented in a different color (e.g., blue). Keeping the spatio-temporal contiguity of stimuli constant, Hughes and colleagues (2020) found that feature transformation (i.e., the transfer of valence) took place for example between two stimuli that shared a color, size, or location. The findings are explained in terms of relational contextual learning (e.g., shared physical features might be conceptualized in a similar manner as relational contextual cues like 'SAME') and inferential reasoning (i.e., currently available propositions). According to this principle the spatio-temporal contiguity of CS and US represent a shared feature of these stimuli. However, the principle predicts that any shared feature of CS and US (e.g., shared background color) might lead to an EC effect. In this sense, Experiment 4 from Chapter 5 aligns relatively closely with the assumptions of the shared feature principles. in this experiment, the same (vs. different) background color might represent a shared feature. Yet, the preregistered pattern was not significant. However, we did find a significant EC effect across conditions. Thus, from a dual-force perspective, this result (even though not expected a priori) might be explicable by adding one additional assumption. If spatio-temporal proximity is the primary factor influencing the encoding of CS and US into a mental episode, then additional bottom-up information should have little impact on the encoding of the two stimuli into a mental episode. One could even argue that any additional piece of information represents a preventing condition on the part of the organism. Yet, this assumption is purely speculative. Still, from a shared feature perspective, a post hoc explanation is more difficult. However, more research is required to make valid arguments for or against each perspective.

7.2.3 The Evolutionary Approach

The findings of Chapter 6 are more readily explicable by a cognitive-ecological approach than by an ecological approach (for a detailed description, see Chapter 3). Yet, as already mentioned in Chapter 3, there are examples of negativity bias in EC (e.g., Levey & Martin, 1975; Baeyens et al., 1990). Additionally, findings from adjacent research disciplines suggest the importance of the evolutionary approach in explaining negativity biases (e.g., decision making: Kahneman & Tversky, 1979; impression formation: Pratto & John, 1991; attribution: Weiner, 1985). Thus, both, the cognitive-ecological and the evolutionary approach might account for the occurrence of valence asymmetries. More interestingly, however, while both approaches are distinguishable at the proximal level (Brunswik, 1955), they might be traced back to the same distal level. For example, both approaches predict attentional advantages (e.g., Unkelbach et al., 2019), and deeper encoding (e.g., Unkelbach et al., 2020).

7.3 Underlying Processes of Valence Asymmetries in Evaluative Conditioning

Although underlying processes of EC were already discussed in Chapter 2, the underlying processes of valence asymmetries have not been addressed yet. Experiment 3 (Chapter 6) failed to provide evidence for a memory explanation due to ceiling effects of participant's performance. However, the post-hoc assessment of memory cannot entirely eliminate the possibility that memory plays a role during CS ratings (Gast, 2018). Yet, other mechanisms might also be able to explain the valence asymmetries presented in Chapter 6. For example, participants might have perceived the valence of rare USs as more extreme and thus, contrasted them away from the frequent USs (e.g., Alves & Imhoff, 2023). Thus, the *isolation effect* (e.g., von Restorff, 1933) as well as three additional mechanisms, which have explanatory potential for the findings of Chapter 6, are elaborated on in more detail below: the *range-frequency theory* (e.g., Parducci, 1965), *figure-ground effects* (e.g., Rubin, 1915), and *habituation* to the frequent US valence (e.g., Mierop et al., 2019).

7.3.1 The Isolation Effect

The isolation effect claims a memory advantage for distinct stimuli (von Restorff, 1933). Distinct information can differ conceptually or physically from its background information (Schmidt & Schmidt, 2017). For example, in a list of pets (background information), a spider represents a conceptually distinct stimulus. Similarly, in a basket of green apples, a red apple constitutes a physically distinct stimulus. Distinct stimuli might be remembered better, because participants organize the isolated and the background items into two separate lists when encoding them (Hunt & Lamb, 2001). According to the incongruity hypothesis, this organization is done in three steps (Schmidt, 1991). The first step takes place during presentation and consists of a comparison between stimulus and active conceptual framework. The attentional response to the stimulus depends on its overlap with the active information. The less overlap, the more attention is drawn to the stimulus. Importantly, this attention depends on the distinctiveness of the stimulus, not on its emotional valence. Increased attention leads to increased storage of information about the stimulus (Eysenck, 1979/2014). In the second step, information about the incongruent event, prior and subsequent events are elaborated, processed and rehearsed. The third phase concerns the memory test. Performance depends on the information that was stored in memory. Thus, strength of memory varies according to the strength of incongruity.

Regarding EC, CS-US presentations that contrast from other CS-US-presentations should have a memory advantage, whereas CS-US pairings that fit into an activated conceptual framework should not have such an advantage. For example, in a list of books written by Jane Austen (e.g., 'Sense and Sensibility', 'Mansfield Park', 'Emma', 'Pride and Prejudice'), a book written by Stephen King (e.g, 'Pet Sematary') would be distinctive, because of the high similarity between the Jane Austen books and the high difference between the Jane Austen books (e.g., female author, love as the main theme, English gentry as protagonists) and the Stephen King book (e.g., male author, horror novel, different era). However, adding three other Stephen King books (e.g., 'The Shining', 'The Green Mile', 'It') removes their uniqueness and thus decreases their incongruity.

The valence memory task in Experiment 3 (Chapter 6) asked participants to indicate the valence of the CS paired with the rare US and a randomly drawn CS that was paired with a frequent US. The identity memory task then asked participants to identify the US that was presented with the displayed CS during conditioning. Thus, in both tasks, participants only needed to recall the 'rare' CS to inform their decision; if the 'rare' CS was not presented, it must have been one of the 'frequent' CSs. Consequently, the memory tasks were unable to distinguish between memory for the 'rare' CS and memory for the 'frequent' CS.

7.3.2 The Range-Frequency Theory

Parducci's (1965) range-frequency theory describes how judgments are made for different categories of stimuli. According to this theory, when a person is given the task of dividing different stimuli into the two categories of 'good' and 'bad', she does so by (a) dividing the psychological range into subranges whose size is independent of the stimulus conditions (i.e., range principle) and (b) assigning the same frequency to each category (i.e., frequency principle). Thus, the range principle states that a range from 'bad' to 'good' is divided in two equal subranges; for example, the 'bad' subrange is mapped on the left side of the center and the 'good' subrange on the right side of the center. The frequency principle ensures that the stimuli are divided equally into the subranges 'good' and 'bad'. Thus, if there are six stimuli, three of them are assigned to the subrange 'good', and the other three to the subrange 'bad'. In this example, range and frequency are consistent with each other.

This correspondence of range and frequency should be present in typical EC experiments. Stimuli are rated on scales that are divided into uniform subranges. These scales can have different numbers of levels (e.g., 5, 7, 9), but the section left to the center usually represents one subrange (e.g., negative), while the section right to the center represents the other subrange (e.g., positive). Moreover, the number of stimuli presented in typical EC experiments does not differ between valences. That is, CSs are paired equally often with positive as with negative USs. Accordingly, the number of judgments for positive and for negative paired CSs does not differ. Thus, the CSs can be evenly allocated to the two subranges. To illustrate the theory, I will use a hypothetical example. In this example, a 5-point Likert-scale ranging from -2 (‘very negative’) to +2 (‘very positive’) is used to evaluate CSs. 6 CS-US-pairings were presented during the conditioning phase, 3 with positive USs, 3 with negative USs. In the measurement phase, the positively paired CSs will be evenly distributed on the positive subrange of the scale (+2, +1, +0), the negatively paired CSs will be judged using the negative subrange of the scale (-2, -1, -0). Thus, there will be an EC effect (i.e., a difference between CS ratings of positively and negatively paired CSs), but more importantly, there will not be a valence asymmetry: the positive mean is 1, the negative mean is -1. Both mean values have the same distance from the center of the scale.

However, there might be cases when range and frequency do not align. For example, inequality in frequencies should lead to a different outcome. Let’s take the example again and assume that there is no equal distribution of positive and negative stimuli, but that positive stimuli are more frequent than negative stimuli (which would mirror the real information ecology). Instead of 3 positive and 3 negative stimuli, we now have 5 positive CS-US pairings and only 1 negative CS-US pairing. In this case, the stimuli cannot be evenly distributed to the subranges. According to Parducci’s (1965) theory, two of the positive stimuli might now be assigned to the negative subrange (three most positive CSs: +2, +1, +0; two least positive CSs: -0, -1). Thus, the mean value of positively paired CSs change to 0.5, while the mean value of the negative CS decreases to -2. That is, the positively paired CSs become (on average) less extreme, while the negatively paired CS becomes more extreme.

However, the range-frequency principle cannot readily explain the findings of Experiment 3 (Chapter 6). In Experiment 3, one of the CSs paired with a frequent US was randomly drawn to be rated. Thus, during measurement phase, only one positive and one negative CS had to be evaluated. Consequently, for the range-frequency theory to explain Experiment 3’s results, one must make the additional assumption that participants mentally distribute the other CSs observed during acquisition onto the Likert scale without actually rating them.

7.3.3 Figure-Ground Effect

Figure-Ground organization has its antecedents in Gestalt psychology: "When an object appears upon a homogeneous field there must be stimulus differentiation (inhomogeneity) in order that the object may be perceived" (Wertheimer, 1938, p. 88). In other words, a figure can only be perceived if it is distinct from its (back)ground. One of the most famous examples is the vase-faces figure by Rubin (1915/1921). In this example, either the faces (white color) or the vase (black color) might be perceived as the figure or the ground of the image. There are a variety of theories that try to explain why which structures should be perceived as the figure or as the ground (for an overview, see Wagemans et al., 2012).

Even though early work on Gestalt principles did not refer to attention (Scholl, 2001), more current research suggests attention as a possible explanation for differences in perception (Wagemans et al., 2012). As differences in salience increase the appearance of figure-ground asymmetries, stimuli assigned as figure (vs. ground stimuli) might 'pop out' with regard to visual features like form, color or brightness (Rothermund & Wentura, 2001). Another such feature might be the valence of the stimuli. Thus, if the rare CS-US pairings represents the figure, the US valence might 'pop out' and thus, lead to stronger evaluation of the respective CS.

US Habituation. The concept of habituation was already introduced in Chapter 1. With regard to the results of Chapter 6, participants might have adapted to the affect of the redundant USs, which reduced the strength of valence transfer to the CS, thus decreasing evaluations of prevalent CSs. Indeed, Mierop and colleagues (2019) showed that US evaluations are sensitive for affective habituation effects. When presented repeatedly (with or without CSs), positive USs became more negative and negative USs became more positive. This effect was stronger for positive than for negative USs. However, this explanation needs the additional assumption that habituation is stronger for frequent USs (e.g., due to valence generalization) than for infrequent USs in a one-to-one stimulus assignment. Still, it might be interesting for future research to implement a control condition to disentangle decreased ratings of frequent CSs due to habituation and increased ratings of infrequent CSs due to distinctiveness.

7.4 Limitations

Although the mechanisms I described in the previous section do have explanatory potential, they have not been tested in this thesis. Additionally, we only tested one assumption of the EvIE, that is the differential frequency of positive and negative information. Thus, even though similarity is more difficult to grasp and highly dependent on context factors (e.g., Koch et al., 2022), it would further support the cognitive-ecological explanation. In a typical EC design, for example, pictures of dogs could be used as similar stimuli and stimuli from other modalities as dissimilar stimuli. It might also be interesting to test the cognitive-ecological explanation and the evolutionary approach against each other; something that I could not provide in this thesis. Still, Hütter and Fiedler (2016) argued that different aspects of the ecology may affect how we learn. If organizing the ecological influences on a continuum regarding their flexibility, evolutionary influences might represent one end of the continuum; these influences are deeply rooted in our biology and evolutionary history and thus rather rigid (cf. preparedness). Other influences, however, might be shaped by personal experiences within a culture; these influences are more flexible. On the very end of the continuum, the ecology is thought to offer influences that are highly flexible, which allows to manipulate and test these influences in a laboratory setting. Evolutionary influences might thus manifest in top-down input (e.g., due to preparedness of specific CS-US combinations, cf. Fiedler & Unkelbach, 2011), while environmental influences can be varied through experimental manipulations (cf. Chapter 6). Thus, using more biologically relevant stimuli might further increase the negativity biases found in Chapter 6's experiments. Benedict and Gast (2019) for example found larger EC effects using fear/disgust- than non-fear/non-disgust-evoking USs. Other researcher, however, found aversive conditioning with biologically significant USs quantitatively comparable to EC (Hermans et al., 2002). Additionally, one might have to argue in which cases conditioning using biologically significant stimuli still accounts for EC. That is, EC usually uses non-aversive stimuli such as flavors or pictures of animals or faces, whereas classical conditioning is characterized by aversive stimuli like food or electric shocks (i.e., USs that are unconditional in a strict sense, as "they elicit a similar and innate/unlearned positive or negative response in all subjects", Hermans et al., 2002, p. 219). Still, an example of EC with biologically significant stimuli might be the conditioning with pictures of animals that are evolutionarily significant (e.g., snakes). The pairing of such USs with CSs, without the simultaneous presence of electric shocks, aligns with a typical EC paradigm. Moreover, negativity bias is attenuated (e.g., Sparks & Ledgerwood, 2019) or even shifts to a positivity bias in older people (e.g., Reed & Carstensen, 2012; for a meta-analysis see Reed et al., 2014). Because this shift is argued to have evolutionary advantages (Carstensen & DeLiema, 2018), it might be interesting to try to replicate Chapter 6's experiments with an older sample. Furthermore, using socially relevant stimuli might expand the findings beyond the scope of non-social stimuli.

Similarly, there are some limitations with regard to the other studies reported in this thesis. First of all, we did not test every path of the dual-force perspective in Chapter 4 (cf. Figure 4.1). Although we provided evidence for the leftmost panel using control conditions, we did not test the influence of weak bottom-up and weak top-down input. However, as mentioned earlier, this combination might be the least interesting one, and it is inherently true even for the extreme cases where spatial and temporal contiguity are absent. Additionally, we did not test the rightmost path of weak bottom-up and strong top-down input. However, evidence for this path might come from other studies (even though this evidence was interpreted in terms of the *how* question). For example, De Houwer (2006) showed that nonwords were evaluated more positively or negatively based on participants' instructions about their co-occurrence with positive or negative USs, even though the stimuli were never actually presented together (see also Moran, Van Dessel et al., 2021). Similarly, Gast and De Houwer (2012) were able to elicit an EC effect in participants linking USs and CSs in an indirect manner. In their study, a positive and a negative US were presented with a number (1 or 2) and a grey square. After the acquisition phase, participants were told that the square covered one specific CS whenever '1' was displayed and another CS whenever '2' was displayed. If the positive US co-occurred with Number 1, the respective CS was rated more positively than the Number 2 CS. Thus, these results show that it is sufficient for information to merely imply the co-occurrence of CS and US to produce an EC effect. Yet, a more direct test of this path would be favorable. Particularly considering the results of Chapter 5, a more thorough examination of the dual-force perspective is necessary. While the experiments from Chapter 4 paint a coherent picture, the results from Chapter 5 are not readily compatible with the dual-force perspective. Instead, the findings of Chapter 5 allow for three possible interpretations: (a) spatio-temporal contiguity holds such significant influence as a bottom-up input that further manipulations contribute little value, (b) the applied manipulations were too subtle, or (c) the manipulations inadvertently affected the CS and US in unanticipated ways. The first possibility might easily be tested by increasing the distance between CS and US and then replicate Experiment 2 (i.e., conceptual fit of CS and US) with adjustments made to the contiguity of CS and US. Additionally, replications of Chapter 5's Experiments 3 and 4 with an adjustment of spatio-temporal contiguity of CS and US (i.e., decrease in proximity) might be an interesting test of the leftmost panel of Figure 4.1. It might be somewhat more challenging to test the orthogonality of bottom-up and top-down input as well as associative and propositional processes. First, one might argue that bottom-up input simply represents plain associations between CS and US. However, spatio-temporal contiguity (bottom-up input) is also discussed to elicit propositions (e.g., "stimuli that co-occur have the same valence", cf. Van Dessel et al., 2018; see also De Houwer, 2018). Thus, a distinction between #3 and #4 (cf. Figure 4.10) seems expedient. Second, top-down input might be equated with propositions. However, it is conceivable that instructions lead to the encoding of CS and US into the same mental episode, while the link that is formed between CS and US is of associative nature. Thus, distinguishing between #1 and #2 is possible. Nevertheless, we did not empirically test these assumptions. Additionally, as emphasized before, a single-process approach to EC might offer many advantages (see Hu et al., 2017a; Hu et al., 2017b). Yet, the results of Chapter 4 are difficult to explain with regard to the *how* question; however, taking the *when* question into account allows for the direct prediction of the data pattern.

To make well-founded statements about occurrence of EC, a more systematic investigation of variables that might influence EC is needed. As previously noted, the number of CS-US pairings might impact EC. However, previous research on this topic has been inconclusive. A more systematic approach to EC would provide insights into how variations in the EC effect are influenced by specific factors, thereby facilitating the answer to both, the *when* and *how* question. Yet, none of the experiments provide such a systematic change of procedural variables.

7.5 Real World Implications

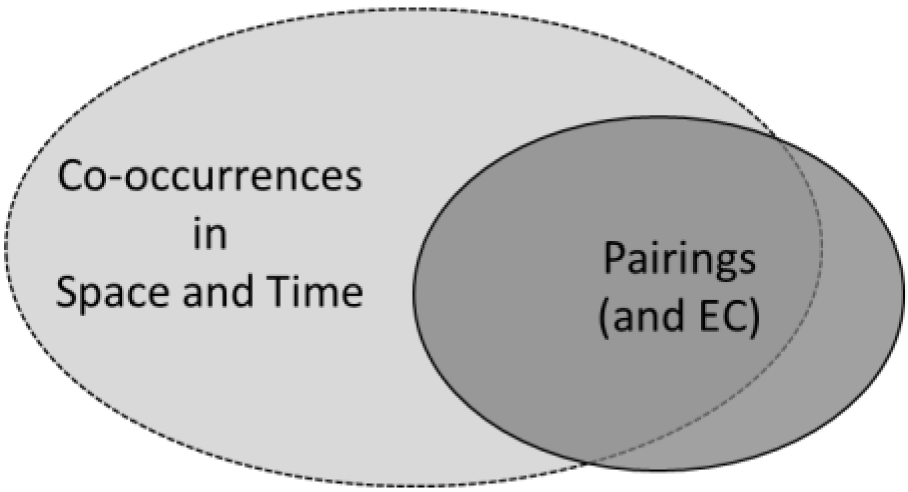
Some researchers point to "the fact that there are many instances in which EC failed to occur" (Walther et al., 2011, p. 17; see also De Houwer et al., 2005). Rozin and colleagues (1998), for example, investigated EC in real world situations. Using shampoo as CS and hair-washing as US, they did not find significant differences between hair-washing-likers and hair-washing-neutrals (Experiment 4). Similarly, they did not find an EC effect when asking participants to smell a neutral scent (CS) during liked or disliked activities (US; Experiment 5). Thus, it was argued that EC might be a rather fragile effect (Jones et al., 2010; see also Field & Moore, 2005). However, the dual-force perspective tested in Chapter 3 allows another explanation for these failures in finding EC effects, especially in the real world. If encoding CS and US into the same mental episode depends on bottom-up and top-down input, it becomes evident that the mere bottom-up input (i.e., the co-occurrence of CS and US) is not sufficient for EC effects to occur. The difference between the real world in terms of the organism's evaluative ecology versus the ecology in EC experiments has been discussed earlier. Yet, it is conceivable that the real world itself can be subdivided into numerous smaller ecologies. This means that it might sometimes be useful to perceive the environment in a broader sense. At other times, however, it might be beneficial to consider only specific aspects of this environment; particularly when only a certain aspect of the environment is currently relevant to the organism. This concept will be further discussed in the next section. Finally, the present findings have some implications for the commercial sector, which are discussed below.

7.5.1 Co-Occurrences and Pairings

Figure 7.2 illustrates the relation between co-occurrences and pairings in a Venn diagram. The light grey area indicates co-occurrences in space and time; that is, spatial and temporal contiguity. The dark circle indicates pairings; that is, encoding CS and US into the same mental episode. The diagram illustrates two points: First, it suggests that the circle of co-occurrences is larger than the circle of pairings. In real life, much fewer situations imply construals that lead from spatial and temporal contiguity (i.e., co-occurrences) to the construal of mental episodes (i.e., pairings). In addition, many factors may prevent the organization of two co-occurring stimuli into the same mental episode (cf. Chapter 4). Second, the pairing circle is not fully part of the co-occurrence circle; there is an area where pairings may form without spatial and temporal contiguity. Differentiating between co-occurrences and pairings might also help to avoid the "tempting mistake […] that EC relies on the dyadic pairing of CS and US and to exclude the domain of instructed and inferred pairings" (Hütter & Fiedler, 2016, p. 10).

Figure 7.2

A Venn Diagram to Illustrate the Relation of Co-Occurrences and Pairings.



Note. The light grey circle indicates temporal and spatial contiguity, and the dark grey circle indicates the perception of pairings. The diagram also illustrates pairings that are not based on stimuli's temporal and spatial contiguity (e.g., 'instructed' EC), and it allows stimuli's spatial and temporal contiguity without becoming a pairing.

For example, De Houwer (2006) simply instructed participants that a positive picture (e.g., of flowers) would always be preceded by two nonwords (e.g., BAYRAM and ENANWAL), while a negative picture (e.g., of mutilated bodies) would be preceded by two other nonwords (e.g., UDIBNON and

SARICIK). Although the pictures were never really presented together, "informing participants that nonwords would be followed by positive or negative photographs was sufficient to change the valence of the nonwords" (De Houwer, 2006, p. 183; see also Gast & De Houwer, 2012, 2013). Yet, even though the dark grey circle includes EC, not all instances in which pairings are formed might lead to (standard) EC effects. However, if propositions (e.g., 'CS hates US') or other processes that take place at a later point in time (e.g., retrieval) do not interfere with the pairing, EC effects should occur.

7.5.2 Learning of Valence Asymmetries

Some instances of negativity biases found in experimental situations might be explained (a) in terms of the EvIE model and (b) by real world phenomena. For example, the finding that monetary losses have a greater impact than wins (e.g., Kahnemann & Tversky, 1984) can be explained by the higher frequency of 'wins' in the real world (i.e., monthly income; Alves et al., 2017b). Similarly, Pratto and John (1991) showed that undesirable personality-traits (e.g., mean, sadistic) attract more attention than desirable personality-traits (e.g., honest, kind). Unkelbach and colleagues (2019) note that "adjectives about personalities in languages allow inferences about the structure of personality" (p. 222), before reviewing studies that show a higher frequency of positive words in language. Thus, in the real world, desirable personality-traits seem to be prevalent, while undesirable personality-traits are rare. Pratto and John's (1991) findings are thus in line with the EvIE model's assumptions. The advantage for negative information then follows an ontogenetic rationale: "children may learn that negative information is rarer, and negative information thereby grabs more attention, is elaborated more, and remembered better" (Unkelbach et al., 2020, p. 146). However, the organism might not always include the 'whole' environment in its evaluation. If an organism plays the lottery, for example, the substance of reality might be narrowed down to wins and losses in this lottery. Because the organism has learned that in lotteries, losses are more prevalent than wins, positive information is rarer and has processing advantages (cf. Chapter 6). Thus, the predictions of the EviE model do not only hold true for the 'whole' social environment, but also apply to more narrowly defined ecologies of the real world.

7.5.3 Advertisement

One real life example that might profit from the findings of this thesis is the advertising sector. Even though “companies have spent enormous amounts of time and energy introducing new brands and products […] between 35% and 45% of product introductions are considered failures (Boulding, Morgan, & Staelin, 1997)” (Walther & Langer, 2008, p. 96). This high number of failures might of course have multiple causes (for an overview, see Walther & Langer, 2008). However, based on the findings of this thesis, one of the reasons for the lack of success might be an advertisement’s design. For example, Lidl’s ‘Du hast die Wahl’ [engl. ‘You have the choice’] campaign displayed two products; the more expensive brand-name product was presented on a black background; the cheaper home-brand product was presented on a white background. Below the ‘You have the choice’ Slogan was for example stated ‘Schmeckt beides gleich gut’ [engl. ‘Tastes both equally well’]. If Lidl tried to transfer the ‘taste-good’ feature (i.e., the valence) of the brand-name products to their home-brand product (while similarly praise their own products lower price), the different background colors might not have been the best approach; at least if assuming that spatio-temporal contiguity is the strongest predictor for EC and any additional information prevents encoding of CS and US into the same mental episode (cf. Chapter 5). In that case, a single-colored background might have been the more promising choice (from a mere EC perspective; of course, other factors are also involved in such ads). However, if the slogan ‘Tastes bot equally well’ is thought of as instruction, this might re-establish the mental episode (cf. Chapter 4).

Another interesting aspect raised by Walther and Langer (2008) is that “Coca Cola does not air commercials on news channels because the company is afraid that the sad feeling induced by bad news may affect the evaluation of their (fun) product” (p. 98). In the light of the mood congruency hypothesis (Bower, 1981), this seems reasonable (but see also Walther & Grigoriadis, 2004). However, based on the findings of Chapter 6, fun commercials represent rare stimuli in the context of frequent bad news and thus, companies like Coca Cola might profit from presenting their fun ads on news channels.

7.6 Conclusion

Although humans are constantly surrounded by stimuli that they interpret as 'good' or 'bad', EC rarely occurs in the real world. One possible explanation for this observation could be that even though CS and US co-occur, they are not perceived as paired. Spatio-temporal contiguity appears to be a crucial factor in encoding stimuli into a shared mental episode. However, this bottom-up input only leads to EC if there is no intervening top-down input from the organism. This dual-force perspective might be useful in explaining *when* an EC effect occurs. Additionally, this perspective proved to be helpful in answering the question of why there is no negativity bias in EC. Because EC experiments consist of an equal amount of positive and negative information, an evolutionary (top-down guided) approach would predict a negativity bias. Thus, from a phylogenetic viewpoint, the lack of negativity bias is surprising. However, adopting a cognitive-ecological (bottom-up guided) approach provides an answer to the question asked. Adjusting the environment of EC experiments to resemble the real world (i.e., more positive than negative stimuli) leads to a negativity bias. The occurrence of a positivity bias in a reversed ecology (i.e., more negative than positive stimuli) further supports the cognitive-ecological approach.

References

Ajzen, I., & Fishbein, M. (2000). Attitudes and the attitude-behavior relation: Reasoned and automatic processes. *European Review of Social Psychology*, *11*(1), 1-33. https://doi.org/10.1080/14792779943000116

Allen, C. T., & Janiszewski, C. A. (1989). Assessing the role of contingency awareness in attitudinal conditioning with implications for advertising research. *Journal of Marketing Research*, *26*(1), 30-43. https://doi.org/10.1177/002224378902600101

Alves, H., Högden, F., Gast, A., Aust, F., & Unkelbach, C. (2020). Attitudes from mere co-occurrences are guided by differentiation. *Journal of Personality and Social Psychology, 119*(3), 560-581. https://doi.org/10.1037/pspa0000193

Alves, H., & Imhoff, R. (2023). Evaluative context and conditioning effects among same and different objects. *Journal of Personality and Social Psychology, 124*(4), 735-753. https://doi.org/10.1037/pspa0000323

Alves, H., Koch, A., & Unkelbach, C. (2016). My friends are all alike—the relation between liking and perceived similarity in person perception. *Journal of Experimental Social Psychology*, *62*, 103-117. https://doi.org/10.1016/j.jesp.2015.10.011

Alves, H., Koch, A., & Unkelbach, C. (2017a). The "common good" phenomenon: Why similarities are positive and differences are negative. *Journal of Experimental Psychology: General, 146*(4), 512-528. https://doi.org/10.1037/xge0000276

Alves, H., Koch, A., & Unkelbach, C. (2017b). Why good is more alike than bad: Processing implications. *Trends in Cognitive Sciences*, *21*(2), 69-79. https://doi.org/10.1016/j.tics.2016.12.006

Alves, H., Koch, A., & Unkelbach, C. (2018a). The differential similarity of positive and negative information–an affect-induced processing outcome?. *Cognition and Emotion, 33*, 1224-1238. https://doi.org/10.1080/02699931.2018.1549022

Alves, H., Koch, A., & Unkelbach, C. (2018b). A cognitive-ecological explanation of intergroup biases.

Psychological Science, 29(7), 1126-1133. https://doi.org/10.1177/09567976187568

Alves, H., Unkelbach, C., Burghardt, J., Koch, A. S., Krüger, T., & Becker, V. D. (2015). A density explanation of valence asymmetries in recognition memory. Memory & Cognition, 43, 896-909. https://doi.org/10.3758/s13421-015-0515-5

Amabile, T. M., & Glazebrook, A. H. (1982). A negativity bias in interpersonal evaluation. *Journal of Experimental Social Psychology, 18*(1), 1-22. https://doi.org/10.1016/0022-1031(82)90078-6

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57*, 1060-1073.

American Psychological Association. (2017). Ethical principles of psychologists and code of conduct (2002, amended effective June 1, 2010, and January 1, 2017). http://www.apa.org/ethics/code/index.html

Anderson, J. R. (2000). *Learning and memory: An integrated approach*. John Wiley & Sons Inc.

Anderson, N. H., & Lampel, A. K. (1965). Effect of context on ratings of personality traits. Psychonomic Science, 3, 433-434. https://doi.org/10.3758/BF03343217

Anisfeld, M., & Lambert, W. E. (1966). When are pleasant words learned faster than unpleasant words?. *Journal of Verbal Learning and Verbal Behavior*, *5*(2), 132-141. https://doi.org/10.1016/S0022-5371(66)80006-3

Aristotle (1999). Nicomachean Ethics (W.D. Ross, Trans.). Batoche Books.

Atthowe, J. M., Jr. (1960). Types of conflict and their resolution: A reinterpretation. *Journal of Experimental Psychology, 59*(1), 1-9. https://doi.org/10.1037/h0046912

Augustine, A. A., Mehl, M. R., & Larsen, R. J. (2011). A positivity bias in written and spoken English and its moderation by personality and gender. *Social Psychological and Personality Science*, *2*(5), 508-515. https://doi.org/10.1177/194855061139915

Axelrod, R. (1980). Effective choice in the prisoner's dilemma. *Journal of Conflict Resolution, 24*(1), 3-25. https://doi.org/10.1177/0022002780024001

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *science*, *211*(4489), 1390-1396. https://doi.org/10.1126/science.7466396

Bache, S. M., & Wickham, H. (2020). magrittr: A Forward-Pipe Operator for R (Version 2.0.1). https://CRAN.R-project.org/package=magrittr

Baeyens, F., Crombez, G., Hendrickx, H., & Eelen, P. (1995). Parameters of human evaluative flavor-flavor conditioning. *Learning and Motivation*, *26*(2), 141-160. https://doi.org/10.1016/0023-9690(95)90002-0

Baeyens, F., Crombez, G., Van den Bergh, O., & Eelen, P. (1988). Once in contact always in contact: Evaluative conditioning is resistant to extinction. *Advances in Behaviour Research and Therapy*, *10*(4), 179-199. https://doi.org/10.1016/0146-6402(88)90014-8

Baeyens, F., & De Houwer, J. (1995). Evaluative conditioning is a qualitatively distinct form of classical conditioning: A reply to Davey (1994). *Behaviour Research and Therapy*, *33*(7), 825-831. https://doi.org/10.1016/0005-7967(95)00021-O

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

Baeyens, F., Eelen, P., Crombez, G., & Van den Bergh, O. (1992). Human evaluative conditioning: Acquisition trials, presentation schedule, evaluative style and contingency awareness. *Behaviour Research and Therapy*, *30*(2), 133-142. https://doi.org/10.1016/0005-7967(92)90136-5

Baeyens, F., Hermans, D., & Eelen, P. (1993). The role of CS-US contingency in human evaluative conditioning. *Behaviour Research and Therapy*, *31*(8), 731-737. https://doi.org/10.1016/0005-7967(93)90003-D

Baeyens, F., Wrzesniewski, A., De Houwer, J., & Eelen, P. (1996). Toilet rooms, body massages, and smells: Two field studies on human evaluative odor conditioning. *Current Psychology, 15*(1), 77-96. https://doi.org/10.1007/BF02686936

Balas, R., & Bar-Anan, Y. (2018). Evaluative conditioning – Theoretical accounts [special issue]. Social *Psychological Bulletin, 13*(3).

Balas, R., & Pochwatko, G. (2021). Learning affective responses through evaluative conditioning - New developments in affective computing. In C. Biele, J. Kacprzyk, W. Kopéc, J. W. Owsiński, A. Romanowski, M. Sikorski, *Digital interaction and machine intelligence* (pp. 259-266). Springer. https://doi.org/10.1007/978-3-031-11432-8_26

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215. https://doi.org/10.1037/0033-295X.84.2.191

Bar-Anan, Y., & Balas, R. (2018). Why does co-occurrence change evaluation? Introduction to a special issue on evaluative conditioning. *Social Psychological Bulletin, 13*(3), 1-10. https://doi.org/10.5964/spb.v13i3.29154

Bar-Anan, Y., & Dahan, N. (2013). The effect of comparative context on evaluative conditioning. *Cognition and Emotion*, *27*(2), 367-375. https://doi.org/10.1080/02699931.2012.712946

Bar-Anan, Y., De Houwer, J., & Nosek, B. A. (2010). Evaluative conditioning and conscious knowledge of contingencies: A correlational investigation with large samples. *Quarterly Journal of Experimental Psychology*, *63*(12), 2313-2335. https://doi.org/10.1080/17470211003802

Bar-Anan, Y., & Nosek, B. A. (2014). A comparative investigation of seven indirect attitude measures. *Behavior Research Methods*, *46*, 668-688. https://doi.org/10.3758/s13428-013-0410-6

Bar-Anan, Y., Nosek, B. A., & Vianello, M. (2009). The sorting paired features task: A measure of association strengths. *Experimental Psychology*, *56*(5), 329-343. https://doi.org/10.1027/1618-3169.56.5.329

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition: Basic processes; Applications* (pp. 1-40). Lawrence Erlbaum.

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology, 62*(6), 893–912. https://doi.org/10.1037/0022-3514.62.6.893

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*(4), 323-370. https://doi.org/10.1037/1089-2680.5.4.323

Baumeister, R. F., & Cairns, K. J. (1992). Repression and self-presentation: When audiences interfere with self-deceptive strategies. *Journal of Personality and Social Psychology, 62*(5), 851-862. https://doi.org/10.1037/0022-3514.62.5.851

Beck, J. (1975). The relation between similarity grouping and perceptual constancy. *American Journal of Psychology, 88*(3), 397-409. https://doi.org/10.2307/1421770

Beck, D. M., & Palmer, S. E. (2002). Top-down influences on perceptual grouping. *Journal of Experimental Psychology: Human Perception and Performance, 28*(5), 1071-1084. https://doi.org/10.1037/0096-1523.28.5.1071

Beckers, T., De Vicq, P., & Baeyens, F. (2009). Evaluative conditioning is insensitive to blocking. *Psychologica Belgica, 49*(1), 41-57. https://doi.org/10.5334/pb-49-1-41

Benedict, T., & Gast, A. (2020). Evaluative conditioning with fear-and disgust-evoking stimuli: no evidence that they increase learning without explicit memory. *Cognition and Emotion*, *34*(1), 42-56. https://doi.org/10.1080/02699931.2019.1646213

Benedict, T., Richter, J., & Gast, A. (2019). The influence of misinformation manipulations on evaluative conditioning. *Acta Psychologica, 194,* 28-36. https://doi.org/10.1016/j.actpsy.2019.01.014

Blask, K., Frings, C., & Walther, E. (2020). CS as an effect: action-based evaluative conditioning depends on temporal contiguity. *Journal of Cognitive Psychology*, *32*(7), 661-667. https://doi.org/10.1080/20445911.2020.1810057

Blask, K., Walther, E., & Frings, C. (2017a). When congruence breeds preference: The influence of selective attention processes on evaluative conditioning. *Cognition and Emotion*, *31*(6), 1127-1139. https://doi.org/10.1080/02699931.2016.1197100

Blask, K., Walther, E., & Frings, C. (2017b). Ignorance reflects preference: The influence of selective ignoring on evaluative conditioning. *Journal of Cognitive Psychology*, *29*(8), 939-948. https://doi.org/10.1080/20445911.2017.1340893

Blask, K., Walther, E., Halbeisen, G., & Weil, R. (2012). At the crossroads: Attention, contingency awareness, and evaluative conditioning. *Learning and Motivation*, *43*(3), 99-106. https://doi.org/10.1016/j.lmot.2012.03.004

Blechert, J., Michael, T., Williams, S. L., Purkis, H. M., & Wilhelm, F. H. (2008). When two paradigms meet: Does evaluative learning extinguish in differential fear conditioning?. *Learning and Motivation*, *39*(1), 58-70. https://doi.org/10.1016/j.lmot.2007.03.003

Bless, H., & Fiedler, K. (2006). Mood and the regulation of information processing and behavior. In J. Forgas (Ed.), *Affect in social thinking and behavior* (pp. 65-84). Psychology Press.

Bless, H., & Schwarz, N. (2010). Mental construal and the emergence of assimilation and contrast effects: The inclusion/exclusion model. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 319-373). Academic Press. https://doi.org/10.1016/S0065-2601(10)42006-7

Bolders, A. C., Band, G. P., & Stallen, P. J. (2012). Evaluative conditioning induces changes in sound valence. *Frontiers in Psychology*, *3*, 106. https://doi.org/10.3389/fpsyg.2012.00106

Borenstein, M. (2009). Effect size for continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 221-235). Russell Sage Foundation.

Bosshard, S., Koller, M., & Walla, P. (2019). Can evaluative conditioning change well-established attitudes towards popular brands? Your brain says yes even though your mouth says no. *Brain Sciences*, *9*(5), 106. https://doi.org/10.3390/brainsci9050106

Boucher, J., & Osgood, C. E. (1969). The pollyanna hypothesis. *Journal of verbal learning and verbal behavior*, *8*(1), 1-8. https://doi.org/10.1016/S0022-5371(69)80002-2

Boulding, W., Morgan, R., & Staelin, R. (1997). Pulling the plug to stop the new product drain. *Journal of Marketing Research*, *34*(1), 164-176. https://doi.org/10.1177/002224379703400l

Bouton, M. E. (2007). *Learning and behavior: A contemporary synthesis*. Sinauer Associates.

Bouton, M. E., & Moody, E. W. (2004). Memory processes in classical conditioning. *Neuroscience & Biobehavioral Reviews*, *28*(7), 663-674. https://doi.org/10.1016/j.neubiorev.2004.09.001

Bower, G. H. (1981). Mood and memory. *American Psychologist, 36*(2), 129-148. https://doi.org/10.1037/0003-066X.36.2.129

Brans, K., Koval, P., Verduyn, P., Lim, Y. L., & Kuppens, P. (2013). The regulation of negative and positive affect in daily life. *Emotion, 13*(5), 926-939. https://doi.org/10.1037/a0032400

Brickman, P., Redfield, J., Harrison, A. A., & Crandall, R. (1972). Drive and predisposition as factors in the attitudinal effects of mere exposure. *Journal of Experimental Social Psychology, 8*(1), 31-44. https://doi.org/10.1016/0022-1031(72)90059-5

Brinkman, L., Todorov, A., & Dotsch, R. (2017). Visualising mental representations: A primer on noise-based reverse correlation in social psychology. *European Review of Social Psychology*, *28*(1), 333-361. https://doi.org/10.1080/10463283.2017.1381469

Brownstein, M., Madva, A., & Gawronski, B. (2019). What do implicit measures measure?. *Wiley Interdisciplinary Reviews: Cognitive Science*, *10*(5), e1501. https://doi.org/10.1002/wcs.1501

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64*(2), 123-152. https://doi.org/10.1037/h0043805

Brunstrom, J. M., & Higgs, S. (2002). Exploring evaluative conditioning using a working memory task. *Learning and Motivation, 33*(4), 433-455. https://doi.org/10.1016/ S0023-9690(02)00007-3

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62*(3), 193-217. https://doi.org/10.1037/h0047470

Carstensen, L. L., & DeLiema, M. (2018). The positivity effect: A negativity bias in youth fades with age. *Current Opinion in Behavioral Sciences*, *19*, 7-12. https://doi.org/10.1016/j.cobeha.2017.07.009

Chan C., Leeper T., Becker J., Schoch D. (2023). rio: A Swiss-army knife for data file I/O (Version 0.5.26). https://cran.r-project.org/package=rio

Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods, 46*(1), 112-130. https://doi.org/10.3758/s13428-013-0365-7

Clark, R. E., Manns, J. R., & Squire, L. R. (2002). Classical conditioning, awareness, and brain systems. *Trends in Cognitive Sciences*, *6*(12), 524-531. https://doi.org/10.1016/S1364-6613(02)02041-7

Clark, R. E., & Squire, L. R. (1998). Classical conditioning and brain systems: the role of awareness. *Science*, *280*(5360), 77-81. https://doi.org/10.1126/science.280.5360.77

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Erlbaum.

Coppens, E., Vansteenwegen, D., Baeyens, F., Vandenbulcke, M., Van Paesschen, W., & Eelen, P. (2006). Evaluative conditioning is intact after unilateral resection of the anterior temporal lobe in humans. *Neuropsychologia*, *44*(5), 840-843. https://doi.org/10.1016/j.neuropsychologia.2005.06.008

Costantini, A. F., & Hoving, K. L. (1973). The effectiveness of reward and punishment contingencies on response inhibition. *Journal of Experimental Child Psychology*, *16*(3), 484-494. https://doi.org/10.1016/0022-0965(73)90009-X

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/10.1177/10888683209113

Corneille, O., & Stahl, C. (2019). 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

Corneille, O., Yzerbyt, V. Y., Pleyers, G., & Mussweiler, T. (2009). Beyond awareness and resources: Evaluative conditioning may be sensitive to processing goals. *Journal of Experimental Social Psychology*, *45*(1), 279-282. https://doi.org/10.1016/j.jesp.2008.08.020

Curtis, V., Aunger, R., & Rabie, T. (2004). Evidence that disgust evolved to protect from risk of disease. *Proceedings of the Royal Society of London, Series B, 271*, 131-133. https://doi.org/10.1098/rsbl.2003.0144

Davey, G. C. (1994). Is evaluative conditioning a qualitatively distinct form of classical conditioning?. *Behaviour Research and Therapy*, *32*(3), 291-299. https://doi.org/10.1016/0005-7967(94)90124-4

Davey, G. C., & McKenna, I. (1983). The effects of postconditioning revaluation of CS1 and UCS following Pavlovian second-order electrodermal conditioning in humans. *The Quarterly Journal of Experimental Psychology Section B, 35*(2), 125-133. https://doi.org/10.1080/14640748308400891

Dawson, M. E., Rissling, A. J., Schell, A. M., & Wilcox, R. (2007). Under what conditions can human affective conditioning occur without contingency awareness? Test of the evaluative conditioning paradigm. *Emotion, 7*(4), 755-766. https://doi.org/10.1037/1528-3542.7.4.755

Dedonder, J., Corneille, O., Yzerbyt, V., & Kuppens, T. (2010). Evaluative conditioning of high-novelty stimuli does not seem to be based on an automatic form of associative learning. *Journal of Experimental Social Psychology, 46*(6), 1118-1121. https://doi.org/10.1016/j.jesp.2010.06.004

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204-211. https://doi.org/10.1016/j.tics.2006.03.007

De Houwer, J. (2006). Using the Implicit Association Test does not rule out an impact of conscious propositional knowledge on evaluative conditioning. *Learning and Motivation*, *37*(2), 176-187. https://doi.org/10.1016/j.lmot.2005.12.002

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. (2011). Evaluative conditioning. In T. R. Schachtman, & S. Reilly (Eds.), *Associative learning and conditioning theory* (pp. 399-417). Oxford University Press.

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

De Houwer, J. (2020). Conditioning is more than association formation: On the different ways in which conditioning research is valuable for clinical psychology. *Collabra: Psychology*, *6*(1), 2. https://doi.org/10.1525/collabra.239

De Houwer, J., Baeyens, F., & Eelen, P. (1994). Verbal evaluative conditioning with undetected US presentations. *Behaviour Research and Therapy*, *32*(6), 629-633. https://doi.org/10.1016/0005-7967(94)90017-5

De Houwer, J., Baeyens, F., & Field, A. P. (2005). Associative learning of likes and dislikes: Some current controversies and possible ways forward. In J. De Houwer, F. Baeyens & A. Field, *Associative learning of likes and dislikes* (pp. 161-174). Psychology Press. https://doi.org/10.4324/9780203024768

De Houwer, J., Barnes-Holmes, D., & Moors, A. (2013). What is learning? On the nature and merits of a functional definition of learning. *Psychonomic Bulletin and Review*, *20*, 631-642. https://doi.org/10.3758/s13423-013-0386-3

De Houwer, J., Baeyens, F., Vansteenwegen, D., & Eelen, P. (2000). Evaluative conditioning in the picture–picture paradigm with random assignment of conditioned stimuli to unconditioned stimuli. *Journal of Experimental Psychology: Animal Behavior Processes, 26*(2), 237-242. https://doi.org/10.1037/0097-7403.26.2.237

De Houwer, J. D., Beckers, T., & Glautier, S. (2002). Outcome and cue properties modulate blocking. *The Quarterly Journal of Experimental Psychology Section A*, *55*(3), 965-985. https://doi.org/10.1080/02724980143000578

De Houwer, J., & Hughes, S. (2016). Evaluative conditioning as a symbolic phenomenon: On the relation between evaluative conditioning, evaluative conditioning via instructions, and persuasion. *Social Cognition*, *34*(5), 480-494. https://doi.org/10.1521/soco.2016.34.5.480

De Houwer, J., & Hughes, S. (2020). Learning to like or dislike: Revealing similarities and differences between evaluative learning effects. *Current Directions in Psychological Science*, *29*(5), 487-491. https://doi.org/10.1177/0963721420924752

De Houwer, J., & Hughes, S. (2023). Learning in individual organisms, genes, machines, and groups: A new way of defining and relating learning in different systems. *Perspectives on Psychological Science*, *18*(3), 649-663. https://doi.org/10.1177/1745691622111488

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin, 135*(3), 347-368. https://doi.org/10.1037/a0014211

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Association learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin, 127*(6), 853-869. https://doi.org/10.1037/0033-2909.127.6.853

De Houwer J., Van Dessel P., Moran T. (2020). Attitudes beyond associations: On the role of propositional representations in stimulus evaluation. *Advances in Experimental Social Psychology, 61*, 127-183. https://doi.org/10.1016/bs.aesp.2019.09.004

Denrell, J. (2005). Why most people disapprove of me: Experience sampling in impression formation. *Psychological Review, 112*(4), 951-978. https://doi.org/10.1037/0033-295X.112.4.951

Dickinson, A. (1980). *Contemporary animal learning theory*. Cambridge University Press.

Dickinson, A. (1994). Instrumental conditioning. In N. J. Mackintosh (Ed.), *Animal learning and cognition* (pp. 45-79). Academic Press. https://doi.org/10.1016/B978-0-08-057169-0.50009-7

Dienes, Z. (2019). How do I know what my theory predicts? *Advances in Methods and Practices in Psychological Science, 2*(4), 364-377. https://doi.org/10.1177/ 2515245919876960

Dijksterhuis, A. (2004). I like myself but I don't know why: Enhancing implicit self-esteem by subliminal evaluative conditioning. Journal of Personality and Social Psychology, 86(2), 345-355. https://doi.org/10.1037/0022-3514.86.2.345

Domjan, M. (2010). *The principles of learning and behavior.* Wadsworth Publishing Company.

Dotsch, R., & Todorov, A. (2012). Reverse correlating social face perception. *Social Psychological and Personality Science*, *3*(5), 562-571. https://doi.org/10.1177/1948550611430272

Doyle, F. L., Klein, L., Kemp, L. J., Moul, C., Richmond, J. L., Eapen, V., Frick, P. J., Kimonis, E. R., Hawes, D. J., Le Pelley, M. E., Mehta, D., & Dadds, M. R. (2022). Learning to like triangles: A longitudinal investigation of evaluative conditioning in infancy. *Developmental Psychobiology*, *64*(3), e22244. https://doi.org/10.1002/dev.22244

Ekman, P., Friesen, W. V., & Ellsworth, P. (1982). What emotion categories or dimensions can observers judge from facial behavior? In P. Ekman (Ed.), *Emotion in the human face* (pp. 39-55). Cambridge University Press.

Eysenck, M. W. (2014). Depth, elaboration, and distinctiveness. In L. S. Cermak, & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 89-118). Lawrence Erlbaum. (Original work published 1979)

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*, 175-191. https://doi.org/10.3758/BF03193146

Fazio, R. H., Eiser, J. R., & Shook, N. J. (2004). Attitude formation through exploration: Valence asymmetries. *Journal of Personality and Social Psychology, 87*(3), 293-311. https://doi.org/10.1037/0022-3514.87.3.293

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69*(6), 1013-1027. https://doi.org/10.1037/0022-3514.69.6.1013

Fiedler, K. (2001). Affective states trigger processes of assimilation and accommodation. In L. L. Martin, & G. L. Clore (Eds.), *Theories of mood and cognition: A user's guidebook* (pp. 85-98). Lawrence Erlbaum.

Fiedler, K., & Bless, H. (2000). The formation of beliefs and the interface of affective and cognitive processes. In N. H. Frijda, A. S. R. Manstead, & S. Bem (Eds*.), Emotions and beliefs: How feelings influence thoughts* (pp. 144–170). Cambridge University Press. https://doi.org/10.1017/CBO9780511659904.006

Fiedler, K., Kutzner, F., & Vogel, T. (2013). Pseudocontingencies: Logically unwarranted but smart inferences. *Current Directions in Psychological Science*, *22*(4), 324-329. https://doi.org/10.1177/09637214134801

Fiedler, K., Schenck, W., Watling, M., & Menges, J. I. (2005). Priming trait inferences through pictures and moving pictures: The impact of open and closed mindsets. Journal of Personality and Social Psychology, 88(2), 229-244. https://doi.org/10.1037/0022-3514.88.2.229

Fiedler, K., & Unkelbach, C. (2011). Evaluative conditioning depends on higher order encoding processes. *Cognition and Emotion, 25*(4), 639-656. https://10.1080/02699931.2010.513497

Field, A. P., & Moore, A. C. (2005). Dissociating the effects of attention and contingency awareness on evaluative conditioning effects in the visual paradigm. *Cognition and Emotion*, *19*(2), 217-243. https://doi.org/10.1080/02699930441000292

Finkenauer, C., & Rimé, B. (1998). Socially shared emotional experiences vs. emotional experiences kept secret: Differential characteristics and consequences. *Journal of Social and Clinical Psychology*, *17*(3), 295-318. https://doi.org/10.1521/jscp.1998.17.3.295

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology, 38*(6), 889-906. https://doi.org/10.1037/0022-3514.38.6.889

Förderer, S., & Unkelbach, C. (2011). Beyond evaluative conditioning! Evidence for transfer of non-evaluative attributes. *Social Psychological and Personality Science*, *2*(5), 479-486. https://doi.org/10.1177/1948550611398413

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

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

Förderer, S., & Unkelbach, C. (2015). Attribute conditioning: Changing attribute-assessments through mere pairings. *Quarterly Journal of Experimental Psychology, 68*(1), 144-164. https://doi.org/10.1080/17470218.2014.939667

Franklin, J. C., Fox, K. R., & Ribeiro, J. D. (2017). How to implement Therapeutic Evaluative Conditioning in a clinical setting. *Journal of Clinical Psychology*, *73*(5), 559-569. https://doi.org/10.1002/jclp.22453

French, A. R., Franz, T. M., Phelan, L. L., & Blaine, B. E. (2013). Reducing Muslim/Arab stereotypes through evaluative conditioning. *The Journal of Social Psychology, 153*(1), 6-9. https://doi.org/10.1080/00224545.2012.706242

Fulcher, E. P., & Cocks, R. P. (1997). Dissociative storage systems in human evaluative conditioning. *Behaviour Research and Therapy*, *35*(1), 1-10. https://doi.org/10.1016/S0005-7967(96)00081-2

Fulcher, E. P., & Hammerl, M. (2001). When all is revealed: A dissociation between evaluative learning and contingency awareness. *Consciousness and Cognition*, *10*(4), 524-549. https://doi.org/10.1006/ccog.2001.0525

Garcia, J., McGowan, B. K., Ervin, F. R., & Koelling, R. A. (1968). Cues: Their relative effectiveness as a function of the reinforcer. *Science, 160*(3829), 794-795. https://doi.org/10.1126/science.160.3829.794

Garcia, J., & Koelling, R. A. (1966). Relation of cue to consequence in avoidance learning. Psychonomic *Science, 4*(1), 123-124. https://doi.org/10.3758/BF03342209

Gaspelin, N., & Luck, S. J. (2018). The role of inhibition in avoiding distraction by salient stimuli. *Trends in Cognitive Sciences, 22*, 79-92. https://doi.org/10.1016/j. tics.2017.11.001

Gast, A. (2018). A declarative memory model of evaluative conditioning. *Social Psychological Bulletin*, *13*(3), 1-23. 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/10.1080/17470218.2012.66506

Gast, A., De Houwer, J., & De Schryver, M. (2012). Evaluative conditioning can be modulated by memory of the CS–US pairings at the time of testing. *Learning and Motivation*, *43*(3), 116-126. https://doi.org/10.1016/j.lmot.2012.06.001

Gast, A., & De Houwer, J. (2013). The influence of extinction and counterconditioning instructions on evaluative conditioning effects. *Learning and Motivation*, *44*(4), 312-325. https://doi.org/10.1016/j.lmot.2013.03.003

Gast, A., Gawronski, B., & De Houwer, J. (2012). Evaluative conditioning: Recent developments and future directions. *Learning and Motivation*, *43*(3), 79-88. https://doi.org/10.1016/j.lmot.2012.06.004

Gast, A., Langer, S., & Sengewald, M. A. (2016). Evaluative conditioning increases with temporal contiguity. The influence of stimulus order and stimulus interval on evaluative conditioning. *Acta Psychologica*, *170*, 177-185. https://doi.org/10.1016/j.actpsy.2016.07.002

Gast, A., & Rothermund, K. (2011a). What you see is what will change: Evaluative conditioning effects depend on a focus on valence. *Cognition and Emotion*, *25*(1), 89-110. https://doi.org/10.1080/02699931003696380

Gast, A., & Rothermund, K. (2011b). I like it because I said that I like it: Evaluative conditioning effects can be based on stimulus-response learning. *Journal of Experimental Psychology: Animal Behavior Processes, 37*(4), 466-476. https://doi.org/10.1037/a0023077

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: Theory, evidence, and open questions. *Advances in Experimental Social Psychology*, *44*, 59-127. https://doi.org/10.1016/B978-0-12-385522-0.00002-0

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). The Guilford Press.

Gawronski, B., & Bodenhausen, G. V. (2018). Evaluative conditioning from the perspective of the associative-propositional evaluation model. *Social Psychological Bulletin*, *13*(3), 1-33. https://doi.org/10.5964/spb.v13i3.28024

Gawronski, B., & Hahn, A. (2018). Implicit measures: Procedures, use, and interpretation. In H. Blanton, J. M. LaCroix, & G. D. Webster (Eds.), *Measurement in social psychology* (pp. 29-55). Routledge.

Gawronski, B., De Houwer, J., & Sherman, J. W. (2020). Twenty-five years of research using implicit measures. *Social Cognition*, 38(Supplement), 1-25. https://doi.org/10.1521/soco.2020.38.supp.s1

Gawronski, B., & Mitchell, D. G. (2014). Simultaneous conditioning of valence and arousal. *Cognition and Emotion*, *28*(4), 577-595. https://doi.org/10.1080/02699931.2013.843506

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., Ye, Y., Rydell, R. J., & De Houwer, J. (2014). Formation, representation, and activation of contextualized attitudes. *Journal of Experimental Social Psychology*, *54*, 188-203. https://doi.org/10.1016/j.jesp.2014.05.010

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., Iskander, M. L., 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

Glashouwer, K. A., Neimeijer, R. A. M., de Koning, M. L., Vestjens, M., & Martijn, C. (2018). Evaluative conditioning as a body image intervention for adolescents with eating disorders. *Journal of Consulting and Clinical Psychology, 86*(12), 1046-1055. https://doi.org/10.1037/ccp0000311

Goldsmith, K., & Dhar, R. (2013). Negativity bias and task motivation: Testing the effectiveness of positively versus negatively framed incentives. *Journal of Experimental Psychology: Applied, 19*(4), 358-366. https://doi.org/10.1037/a0034415

Gottman, J. M., & Krokoff, L. J. (1989). Marital interaction and satisfaction: A longitudinal view. *Journal of Consulting and Clinical Psychology, 57*(1), 47-52. https://doi.org/10.1037/0022-006X.57.1.47

Green, L. J., Luck, C. C., Gawronski, B., & Lipp, O. V. (2021). Contrast effects in backward evaluative conditioning: Exploring effects of affective relief/disappointment versus instructional information. *Emotion, 21*(2), 350-359. https://doi.org/10.1037/ emo0000701

Greenwald, A. G., & Lai, C. K. (2020). Implicit social cognition. *Annual review of psychology*, *71*, 419-445. https://doi.org/10.1146/annurev-psych-010419-050837

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464-1480. https://doi.org/10.1037/0022-3514.74.6.1464

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review, 114*(2), 211-244. https://doi.org/10.1037/0033-295X.114.2.211

Groves, P. M., & Thompson, R. F. (1970). Habituation: a dual-process theory. *Psychological Review*, *77*(5), 419-450. https://doi.org/10.1037/h0029810

Halbeisen, G., Blask, K., Weil, R., & Walther, E. (2014). The role of recollection in evaluative conditioning. *Journal of Experimental Social Psychology, 55*, 162-168. https://doi.org/10.1016/j.jesp.2014.07.005

Halbeisen, G., & Walther, E. (2015). Dual-task interference in evaluative conditioning: Similarity matters! *Quarterly Journal of Experimental Psychology, 68*(10), 2008-2021. https://doi.org/10.1080/17470218.2014.1002506

Halbeisen, G., & Walther, E. (2016). Evaluative conditioning is sensitive to the encoding of CS-US contingencies. *Social Cognition*, *34*(5), 462-479. https://doi.org/10.1521/soco.2016.34.5.462

Halbeisen, G., Walther, E., & Schneider, M. (2017). Evaluative conditioning and the development of attitudes in early childhood. *Child Development*, *88*(5), 1536-1543. https://doi.org/10.1111/cdev.12657

Hammerl, M., Bloch, M., & Silverthorne, C. P. (1997). Effects of US-alone presentations on human evaluative conditioning. *Learning and Motivation*, *28*(4), 491-509. https://doi.org/10.1006/lmot.1997.0970

Hammerl, M., & Grabitz, H. J. (1993). Human evaluative conditioning: Order of stimulus presentation. *Integrative Physiological and Behavioral Science*, *28*, 191-194. https://doi.org/10.1007/BF02691227

Hammerl, M., & Grabitz, H. J. (1996). Human evaluative conditioning without experiencing a valued event. *Learning and Motivation*, *27*(3), 278-293. https://doi.org/10.1006/lmot.1996.0015

Hammerl, M., & Grabitz, H. J. (2000). Affective-evaluative learning in humans: A form of associative learning or only an artifact?. *Learning and Motivation*, *31*(4), 345-363. https://doi.org/10.1006/lmot.2000.1059

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology, 54*(6), 917-924. https://doi.org/10.1037/0022-3514.54.6.917

Hasford, J., Kidwell, B., & Hardesty, D. M. (2018). Emotional ability and associative learning: How experiencing and reasoning about emotions impacts evaluative conditioning. *Journal of Consumer Research*, *45*(4), 743-760. https://doi.org/10.1093/jcr/ucy026

Hawkins, R. D. (1989). A simple circuit model for higher-order features of classical conditioning. In J. H. Byrne & W. O. Bell. (Eds.), *Neural models of plasticity* (pp. 73-93). Academic Press.

Hermans, D., Vansteenwegen, D., Crombez, G., Baeyens, F., & Eelen, P. (2002). Expectancy-learning and evaluative learning in human classical conditioning: Affective priming as an indirect and unobtrusive measure of conditioned stimulus valence. *Behaviour Research and Therapy*, *40*(3), 217-234. https://doi.org/10.1016/S0005-7967(01)00006-7

Heycke, T., Aust, F., & Stahl, C. (2017). Subliminal influence on preferences? A test of evaluative conditioning for brief visual conditioned stimuli using auditory unconditioned stimuli. *Royal Society Open Science*, *4*(9), 160935. https://doi.org/10.1098/rsos.160935

Heycke, T., & Stahl, C. (2020). No evaluative conditioning effects with briefly presented stimuli. *Psychological Research*, *84*(4), 1020-1027. https://doi.org/10.1007/s00426-018-1109-1

Hilbig, B. E. (2009). Sad, thus true: Negativity bias in judgments of truth. *Journal of Experimental Social Psychology, 45*(4), 983-986. https://doi.org/10.1016/j.jesp.2009.04.012

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. (2004). Event files: Feature binding in and across perception and action. *Trends in cognitive sciences*, *8*(11), 494-500. https://doi.org/10.1016/j.tics.2004.08.007

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), 849-878. https://doi.org/10.1017/S0140525X01000103

Hommel, B., & Stevenson, N. (2021). An event-coding account of attitudes. Psychonomic *Bulletin & Review, 28*(6), 2057-2063. https://doi.org/10.3758/s13423-021-01969-y

Horowitz, F. D. (1992). John B. Watson's legacy: Learning and environment. *Developmental Psychology*, *28*(3), 360-367. https://doi.org/10.1037/0012-1649.28.3.360

Houben, K., Havermans, R. C., & Wiers, R. W. (2010). Learning to dislike alcohol: Conditioning negative implicit attitudes toward alcohol and its effect on drinking behavior. *Psychopharmacology*, *211*, 79-86. https://doi.org/10.1007/s00213-010-1872-1

Högden, F., Hütter, M., & Unkelbach, C. (2018). Does evaluative conditioning depend on awareness? Evidence from a continuous flash suppression paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(10), 1641-1657. https://doi.org/10.1037/xlm0000533

Högden, F., & Unkelbach, C. (2021). The role of relational qualifiers in attribute conditioning: Does disliking an athletic person make you unathletic? *Personality and Social Psychology Bulletin, 47*(4), 643-656. https://doi.org/10.1177/ 0146167220945538

Hu, X., Gawronski, B., & Balas, R. (2017a). Propositional versus dual-process accounts of evaluative conditioning: I. the effects of co-occurrence and relational information on implicit and explicit evaluations. *Personality and Social Psychology Bulletin, 43*(1), 17-32. https://doi.org/10.1177/0146167216673351

Hu, X., Gawronski, B., & Balas, R. (2017b). Propositional versus dual-process accounts of evaluative conditioning: II. The effectiveness of counter-conditioning and counterinstructions in changing implicit and explicit evaluations. *Social Psychological and Personality Science, 8*(8), 858-866. https://doi.org/10.1177/1948550617691094

Hugdahl, K., & Kärker, A. C. (1981). Biological vs experiential factors in phobic conditioning. *Behaviour Research and Therapy*, *19*(2), 109-115. https://doi.org/10.1016/0005-7967(81)90034-6

Hughes, S., De Houwer, J., & Barnes-Holmes, D. (2016). The moderating impact of distal regularities on the effect of stimulus pairings. *Experimental Psychology*. https://doi.org/10.1027/1618-3169/a000310

Hughes, S., Barnes-Holmes, D., Van Dessel, P., de Almeida, J. H., Stewart, I., & De Houwer, J. (2018). On the symbolic generalization of likes and dislikes. *Journal of Experimental Social Psychology*, *79*, 365-377. https://doi.org/10.1016/j.jesp.2018.09.002

Hughes, S., De Houwer, J., Mattavelli, S., & Hussey, I. (2020). The shared features principle: If two objects share a feature, people assume those objects also share other features. *Journal of Experimental Psychology: General, 149*(12), 2264-2288. https://doi.org/10.1037/xge0000777

Hughes, S., Mattavelli, S., & De Houwer, J. (2018). Examining the impact of distance as a contextual cue in evaluative conditioning. *PloS One*, *13*(10), e0204855. https://doi.org/10.1371/journal.pone.0204855

Hughes, S., Ye, Y., & De Houwer, J. (2019). Evaluative conditioning effects are modulated by the nature of contextual pairings. *Cognition and Emotion*, *33*(5), 871-884. https://doi.org/10.1080/02699931.2018.1500882

Hughes, S., Ye, Y., van Dessel, P., & De Houwer, J. (2019). When people co-occur with good or bad events: Graded effects of relational qualifiers on evaluative conditioning. *Personality and Social Psychology Bulletin*, *45*(2), 196-208. https://doi.org/10.1177/0146167218781340

Humphrey, G. (2013). *The nature of learning: In its relation to the living system*. Routledge. (Original work published 1933)

Hunt, R. R., & Lamb, C. A. (2001). What causes the isolation effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(6), 1359-1366. https://doi.org/10.1037/0278-7393.27.6.1359

Hütter, M. (2022). An integrative review of dual-and single-process accounts of evaluative conditioning. *Nature Reviews Psychology*, *1*(11), 640-653. https://doi.org/10.1038/s44159-022-00102-7

Hütter, M., & Fiedler, K. (2016). Conceptual, theoretical, and methodological challenges in evaluative conditioning research. *Social Cognition*, *34*(5), 343-356. https://doi.org/10.1521/soco.2016.34.5.343

Hütter, M., Kutzner, F., & Fiedler, K. (2014). What is learned from repeated pairings? On the scope and generalizability of evaluative conditioning. *Journal of Experimental Psychology: General, 143*(2), 631-643. https://doi.org/10.1037/a0033409

Hütter, M., & Sweldens, S. (2013). Implicit misattribution of evaluative responses: Contingency-unaware evaluative conditioning requires simultaneous stimulus presentations. *Journal of Experimental Psychology: General, 142*(3), 638-643. https://doi.org/10.1037/a0029989

Hütter, M., Sweldens, S., Stahl, C., Unkelbach, C., & Klauer, K. C. (2012). Dissociating contingency awareness and conditioned attitudes: Evidence of contingency-unaware evaluative conditioning. *Journal of Experimental Psychology: General, 141*(3), 539-557. https://doi.org/10.1037/a0026477

Hütter, M., & Tigges, D. (2019). On the external validity of evaluative conditioning: Evaluative responses generalize to modified instances of conditioned stimuli. *Journal of Experimental Social Psychology*, *84*, 103824. https://doi.org/10.1016/j.jesp.2019.103824

Ingendahl, M., & Vogel, T. (2022). Stimulus evaluation in the eye of the beholder: Big five personality traits explain variance in normed picture sets. *Personality Science*, *3*, 1-21. https://doi.org/10.5964/ps.7951

Ingendahl, M., & Vogel, T. (2023). (Why) Do Big Five Personality Traits Moderate Evaluative Conditioning? The Role of US Extremity and Pairing Memory. *Collabra: Psychology*, *9*(1), 74812. https://doi.org/10.1525/collabra.74812

Ingendahl, M., Vogel, T., Maedche, A., & Wänke, M. (2023). Brand placements in video games: How local in-game experiences influence brand attitudes. *Psychology & Marketing, 40*(2), 274-287. https://doi.org/ 10.1002/mar.21770

Ingendahl, M., Woitzel, J., & Alves, H. (2023). Just Playing the Role of Good Study Participants? Evaluative Conditioning, Demand Compliance, and Agreeableness. *Social Psychological and Personality Science*, 19485506231198653. https://doi.org/10.1177/1948550623119865

Ingendahl, M., Woitzel, J., Propheter, N., Wänke, M., & Alves, H. (2023c). From deviant likes to reversed effects: Re-investigating the contribution of unaware evaluative conditioning to attitude formation. *Collabra: Psychology*, *9*(1). https://doi.org/10.1525/collabra.87462

Ioannou, A., & Anastassiou-Hadjicharalambous, X. (2021). Non-associative learning. *Encyclopedia of evolutionary psychological science*, 5419-5432.

Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology, 75*(4), 887-900. https://doi.org/10.1037/0022-3514.75.4.887

James, W. (1884). What is an emotion? *Mind, 9*, 188-205.

Jones, C. R., Fazio, R. H., & Olson, M. A. (2009). Implicit misattribution as a mechanism underlying evaluative conditioning. *Journal of Personality and Social Psychology*, *96*(5), 933. https://doi.org/10.1037/a0014747

Jones, C. R., Olson, M. A., & Fazio, R. H. (2010). Evaluative conditioning: The “how” question. *Advances in Experimental Social Psychology*, *43,* 205-255. https://doi.org/10.1016/S0065-2601(10)43005-1

Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, *47*(2), 263-292. https://doi.org/10.2307/1914185

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39*(4), 341-350. https://doi.org/10.1037/0003-066X.39.4.341

Kamin, L. J. (1968). ''Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: stimulation*. University of Miami Press.

Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior.* Appleton-Century-Crofts.

Kanouse, D. E., & Hanson, L. R., Jr. (1972). Negativity in evaluations. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 47-62). General Learning Press.

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

Kattner, F. (2012). Revisiting the relation between contingency awareness and attention: evaluative conditioning relies on a contingency focus. *Cognition and Emotion*, *26*(1), 166-175. https://doi.org/10.1080/02699931.2011.565036

Kattner, F. (2014). Reconsidering the (in) sensitivity of evaluative conditioning to reinforcement density and CS–US contingency. *Learning and Motivation, 45,* 15-29. https://doi.org/10.1016/j.lmot.2013.09.002

Kattner, F., & Green, C. S. (2015). Cue competition in evaluative conditioning as a function of the learning process. *Acta Psychologica*, *162*, 40-50. https://doi.org/10.1016/j.actpsy.2015.09.013

Kattner, F., Ellermeier, W., & Tavakoli, P. (2012). Both trace and delay conditioning of evaluative responses depend on contingency awareness. *Learning and Motivation, 43*(1-2), 35-47. https://doi.org/10.1016/j.lmot.2012.01.004

Kim, J. C., Sweldens, S., & Hütter, M. (2016). The symmetric nature of evaluative memory associations: Equal effectiveness of forward versus backward evaluative conditioning. *Social Psychological and Personality Science*, *7*(1), 61-68. https://doi.org/10.1177/19485506155992

Kitayama, S., & Karasawa, M. (1997). Implicit self-esteem in Japan: Name letters and birthday numbers. *Personality and Social Psychology Bulletin*, *23*(7), 736-742. https://doi.org/10.1177/0146167297237006

Klauer, K. C., & Musch, J. (1999). Eine normierung unterschiedlicher aspekte der evaluativen bewertung von 92 substantiven [A standardization of various aspects of the evaluation of 92 nouns]. *Zeitschrift für Sozialpsychologie, 30*(1), 1-11. https://doi.org/10.1024/0044-3514.30.1.1

Koch, A., Alves, H., Krüger, T., & Unkelbach, C. (2016). A general valence asymmetry in similarity: Good is more alike than bad. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(8), 1171-1192. https://doi.org/10.1037/xlm0000243

Koch, A., Speckmann, F., & Unkelbach, C. (2022). Q-SpAM: How to efficiently measure similarity in online research. *Sociological Methods & Research*, *51*(3), 1442-1464. https://doi.org/10.1177/004912412091493

Koch, I., Stahl, C. (2017). Lernen – Assoziationsbildung, Konditionierung und implizites Lernen. In J. Müsseler & M. Rieger (Eds.), *Allgemeine Psychologie* (pp. 319-355). Springer. https://doi.org/10.1007/978-3-642-53898-8_10

Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin, 110*(3), 499-519. https://doi.org/10.1037/0033-2909.110.3.499

Kuchinke, L., Fritsch, N., & Müller, C. J. (2015). Evaluative conditioning of positive and negative valence affects P1 and N1 in verbal processing. *Brain Research*, *1624*, 405-413. https://doi.org/10.1016/j.brainres.2015.07.059

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

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*(3), 480-498. https://doi.org/10.1037/0033-2909.108.3.480

Kurdi, B., & Banaji, M. R. (2019). Attitude change via repeated evaluative pairings versus evaluative statements: Shared and unique features. *Journal of Personality and Social Psychology, 116*(5), 681-703. https://doi.org/10.1037/pspa0000151

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

Kurdi, B., Ratliff, K. A., & Cunningham, W. A. (2021). Can the Implicit Association Test serve as a valid measure of automatic cognition? A response to Schimmack (2021). *Perspectives on Psychological Science*, *16*(2), 422-434. https://doi.org/10.1177/1745691620904080

Laane, K., Aru, J., & Dickinson, A. (2010). Non-competitive liking for brands. No blocking in evaluative conditioning. *Appetite, 54*(1), 100-107. https://doi.org/ 10.1016/j.appet.2009.09.012

Lachman, S. J. (1997). Learning is a process: Toward an improved definition of learning. *The Journal of Psychology*, *131*(5), 477-480. https://doi.org/10.1080/00223989709603535

Lakens, D. (2014). Performing high-powered studies efficiently with sequential analyses. *European Journal of Social Psychology, 44*(7), 701-710. https://doi.org/10.1002/ejsp.2023

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). International Affective Picture System (IAPS): Instruction manual and affective ratings, Technical Report A-8. Gainesville: The Center for Research in Psychophysiology, University of Florida.

Lany, J., & Saffran, J. R. (2013). Statistical learning mechanisms in infancy. In Rubenstein, J. & Rakic, P. (Eds.), Comprehensive developmental neuroscience: Neural circuit development and function in the brain (pp. 231-248). Academic Press.

Leising, D., Scherbaum, S., Locke, K. D., & Zimmermann, J. (2015). A model of "substance" and "evaluation" in person judgments. *Journal of Research in Personality*, *57*, 61-71. https://doi.org/10.1016/j.jrp.2015.04.002

Levey, A. B., & Martin, I. (1975). Classical conditioning of human 'evaluative' responses. *Behaviour Research and Therapy*, *13*(4), 221-226. https://doi.org/10.1016/0005-7967(75)90026-1

Levey, A. B., & Martin, I. (1987). Evaluative conditioning a case for hedonic transfer. In H. J. Eysenck, & I. Martin (Eds.), *Theoretical foundations of Behavior Therapy* (pp. 113-131). Springer.

Lewicki, P., Hill, T., & Czyzewska, M. (1992). Nonconscious acquisition of information. American Psychologist, 47(6), 796-801. https://doi.org/10.1037/0003-066X.47.6.796

Lewin, K. (1936). Principles of topological psychology. (F. Heider & G. M. Heider, Trans.). McGraw-Hill. https://doi.org/10.1037/10019-000

Lipp, O. V., & Purkis, H. M. (2005). No support for dual process accounts of human affective learning in simple Pavlovian conditioning. *Cognition and Emotion*, *19*(2), 269-282. https://doi.org/10.1080/02699930441000319

Lipsey, M. W., & Wilson, D. B. (2001). Practical meta-analysis. Sage Publications.

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

Luck, C. C., & Lipp, O. V. (2020). Relapse of evaluative learning—Evidence for reinstatement, renewal, but not spontaneous recovery, of extinguished evaluative learning in a picture–picture evaluative conditioning paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(6), 1178-1206. https://doi.org/10.1037/xlm0000785

Mallan, K. M., Lipp, O. V., & Libera, M. (2008). Affect, attention, or anticipatory arousal? Human blink startle modulation in forward and backward affective conditioning. *International Journal of Psychophysiology*, *69*(1), 9-17. https://doi.org/10.1016/j.ijpsycho.2008.02.005

Manis, M., Cornell, S. D., & Moore, J. C. (1974). Transmission of attitude relevant information through a communication chain. *Journal of Personality and Social Psychology, 30*(1), 81-94. https://doi.org/10.1037/h0036639

March, D. S., Olson, M. A., & Fazio, R. H. (2018). The implicit misattribution model of evaluative conditioning. *Social Psychological Bulletin, 13*(3), 1-25. https://doi.org/10.5964/spb.v13i3.27574

Martin, I., & Levey, A. B. (1978). Evaluative conditioning. *Advances in Behaviour research and Therapy*, *1*(2), 57-101. https://doi.org/10.1016/0146-6402(78)90013-9

Martin, I., & Levey, A. (1994). The evaluative response: Primitive but necessary. *Behaviour Research and Therapy*, *32*(3), 301-305. https://doi.org/10.1016/0005-7967(94)90125-2

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*(2), 314-324. https://doi.org/10.3758/s13428-011-0168-7

Matlin, M. W., & Stang, D. J. (1978). *The Pollyanna principle: Selectivity in language, memory, and thought.* Schenkman Publishing Company.

Mazur, J. E. & Odum, A. L. (2023). *Learning and Behavior*. Routledge.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. Psychological Review, 100(2), 254-278. https://doi.org/10.1037/0033-295X.100.2.254

Meyer, W. J., & Offenbach, S. I. (1962). Effectiveness of reward and punishment as a function of task complexity. *Journal of Comparative and Physiological Psychology, 55*(4), 532-534. https://doi.org/10.1037/h0049119

Mierop, A., Bret, A., Yzerbyt, V., Dumas, R., & Corneille, O. (2019). Evaluative influences of CS-US pairings are non-reciprocal. *Social Psychological Bulletin*, *14*(1), 1-16. https://doi.org/10.32872/spb.v14i1.31408

Mierop, A., Hütter, M., & Corneille, O. (2017). Resource availability and explicit memory largely determine evaluative conditioning effects in a paradigm claimed to be conducive to implicit attitude acquisition. *Social Psychological and Personality Science, 8*(7), 758-767. https://doi.org/10.1177/1948550616687093

Mierop, A., Maurage, P., & Corneille, O. (2020). Cognitive load impairs evaluative conditioning, even when individual CS and US stimuli are successfully encoded. *International Review of Social Psychology, 33*(1). https://doi.org/10.5334/irsp.339

Mitchell, C. J., Anderson, N. E., & Lovibond, P. F. (2003). Measuring evaluative conditioning using the Implicit Association Test. *Learning and Motivation*, *34*(2), 203-217. https://doi.org/10.1016/S0023-9690(03)00003-1

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

Mitchell, C. J., & Lovibond, P. F. (2002). Backward and forward blocking in human electrodermal conditioning: Blocking requires an assumption of outcome additivity. *The Quarterly Journal of Experimental Psychology Section B*, *55*(4), 311-329. https://doi.org/10.1080/0272499024400002

Moors, A., & De Houwer, J. (2006). Automaticity: A Theoretical and Conceptual Analysis. *Psychological Bulletin, 132*(2), 297–326. https://doi.org/10.1037/0033-2909.132.2.297

Moran, T. (2024). The effect of irrelevant pairings on evaluative responses. *Journal of Experimental Social Psychology, 112*, 104602. https://doi.org/10.1016/j.jesp.2024.104602

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

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., & Nosek, B. A. (2017). The effect of the validity of co-occurrence on automatic and deliberate evaluations. *European Journal of Social Psychology*, *47*(6), 708-723. https://doi.org/10.1002/ejsp.2266

Moran, T., Dror, T., & Bar-Anan, Y. (2020). Testing the judgment-related account for the extinction of evaluative conditioning. *Cognition and Emotion*, *34*(8), 1690-1703. https://doi.org/10.1080/02699931.2020.1798878

Moran, T., Hughes, S., Hussey, I., Vadillo, M. A., Olson, M. A., Aust, F., Bading, K., Balas, R., Benedict, T., Corneille, O., Douglas, S. B., Ferguson, M. J., Fritzlen, K. A., Gast, A., Gawronski, B., Giménez-Fernández, T., Hanusz, K., Heycke, T., Högden, F., …, & De Houwer, J. (2021). Incidental attitude formation via the surveillance task: A preregistered replication of the Olson and Fazio (2001) study. *Psychological Science*, *32*(1), 120-131. https://doi.org/10.1177/0956797620968

Moran, T., Nudler, Y., & Bar-Anan, Y. (2023). Evaluative conditioning: Past, present, and future. *Annual Review of Psychology*, *74*, 245-269. https://doi.org/10.1146/annurev-psych-032420-031815

Moran, T., Van Dessel, P., Smith, C. T., & De Houwer, J. (2021). Can (instructions about) stimulus pairings influence automatic and self-reported evaluations in the presence of more diagnostic evaluative information?. *Personality and Social Psychology Bulletin*, *47*(8), 1249-1263. https://doi.org/10.1177/014616722096463

Morey, R. D., Rouder, J. N., Jamil, T., & Morey, M. R. D. (2015). Package 'BayesFactor': Computation of Bayes Factors for Common Designs (R Package Version 09.12–4.2) [Computer Software]. Retrieved from https:// cran.r-project.org/web/package s/BayesFactor/BayesFactor.pdf

Mowrer, O. H. (1960). *Learning theory and behavior.* Wiley.

Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review, 110*(3), 472. https://doi.org/10.1037/0033- 295X.110.3.472

Myers, J. L., Reilly, R. E., & Taub, H. A. (1961). Differential cost, gain, and relative frequency of reward in a sequential choice situation. *Journal of Experimental Psychology, 62*(4), 357-360. https://doi.org/10.1037/h0047103

Nasrallah, M., Carmel, D., & Lavie, N. (2009). Murder, she wrote: Enhanced sensitivity to negative word valence. *Emotion, 9*(5), 609-618. https://doi.org/10.1037/a0016305

Navon, M., & Bar-Anan, Y. (2023). Is the automatic evaluation of individual group members inherently biased by their group membership?. *Journal of Experimental Social Psychology, 108*, 104491. https://doi.org/10.1016/j.jesp.2023.104491

Necka, E. A., Cacioppo, S., Norman, G. J., & Cacioppo, J. T. (2016). Measuring the prevalence of problematic respondent behaviors among MTurk, campus, and community participants. *PLoS One, 11*(6), Article e0157732. https://doi.org/ 10.1371/journal.pone.0157732

Neisser, U. (2014). Cognitive psychology (Classic ed.). Psychology Press. (Original work published 1967)

Nelson, J. P., & Kennedy, P. E. (2009). The use (and abuse) of meta-analysis in environmental and natural resource economics: an assessment. *Environmental and Resource Economics, 42,* 345-377. https://doi.org/10.1007/s10640-008-9253-5

Newman, A., Bavik, Y. L., Mount, M., & Shao, B. (2021). Data collection via online platforms: Challenges and recommendations for future research. *Applied Psychology, 70*(3), 1380-1402. https://doi.org/10.1111/apps.12302

Nuttin, J. M. (1985). Narcissism beyond Gestalt and awareness: The name letter effect. *European Journal of Social Psychology*, *15*(3), 353-361. https://doi.org/10.1002/ejsp.2420150309

Nuttin, J. M. (1987). Affective consequences of mere ownership: The name letter effect in twelve European languages. *European journal of social psychology*, *17*(4), 381-402. https://doi.org/10.1002/ejsp.2420170402

Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological science*, *12*(5), 413-417. https://doi.org/10.1111/1467-9280.00376

Olson, M. A., & Fazio, R. H. (2006). Reducing automatically activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin*, *32*(4), 421-433.

Öhman, A., Eriksson, A., & Olofsson, C. (1975b). One-Trial Learning and Superior Resistance to Extinction of Autonomic Responses Conditioned to Potentially Phobic Stimuli. *Journal of Comparative and Physiological Psychology, 88*(2), 619-627. https://doi.org/10.1037/h0078388

Öhman, A., Erixon, G., & Löfberg, I. (1975a). Phobias and preparedness: Phobic versus neutral pictures as conditioned stimuli for human autonomic responses. *Journal of Abnormal Psychology, 84*(1), 41-45. https://doi.org/10.1037/h0076255

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology, 80*(3), 381-396. https://doi.org/10.1037/0022-3514.80.3.381

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108*(3), 483-522. https://doi.org/10.1037/0033-295X.108.3.483

Özgen, E. (2004). Language, learning, and color perception. *Current Directions in Psychological Science, 13*(3), 95-98. https://doi.org/10.1111/j.0963-7214.2004.00282

Palmer, S. E. (1992). Common region: A new principle of perceptual grouping. *Cognitive Psychology, 24*(3), 436-447. https://doi.org/10.1016/0010-0285(92)90014-S

Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review, 72*(6), 407-418. https://doi.org/10.1037/h0022602

Pavlov, I. P. (2010). Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. *Annals of Neurosciences, 17*(3), 136-141. (Original work published 1927)

Pearce, J. M., & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology, 52*(1), 111-139. https://doi.org/10.1146/annurev.psych.52.1.111

Peeters, G. (1971). The positive-negative asymmetry: On cognitive consistency and positivity bias. European Journal of Social Psychology, 1, 455-474. https://doi.org/10.1002/ejsp.2420010405

Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European Review of Social Psychology, 1*, 33-60. https://doi.org/10.1080/14792779108401856

Penney, R. K., & Lupton, A. A. (1961). Children's discrimination learning as a function of reward and punishment. *Journal of Comparative and Physiological Psychology, 54*(4), 449-451. https://doi.org/10.1037/h0045445

Piaget, J. (1954). The construction of reality in the child. Basic Books. https://doi.org/10.1037/11168-000

Pleyers, G., Corneille, O., Yzerbyt, V., & Luminet, O. (2009). Evaluative conditioning may incur attentional costs. *Journal of Experimental Psychology: Animal Behavior Processes, 35*(2), 279–285. https://doi.org/10.1037/a0013429

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, Research, and Experience* (Vol. 1. Theories of Emotion, pp. 3-31). Academic Press. https://doi.org/10.3389/fpsyg.2012.00339

Poulton, R., & Menzies, R. G. (2002). Non-associative fear acquisition: A review of the evidence from retrospective and longitudinal research. *Behaviour Research and Therapy, 40*(2), 127-149. https://doi.org/10.1016/S0005-7967(01)00045-6

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology, 61*(3), 380-391. https://doi.org/10.1037/0022-3514.61.3.380

Prescott, J., & Murphy, S. (2009). Inhibition of evaluative and perceptual odour–taste learning by attention to the stimulus elements. *Quarterly Journal of Experimental Psychology, 62*(11), 2133-2140. https://doi.org/10.1080/17470210903031169

Qualtrics, Provo, UT. https://www.qualtrics.com

Radvansky, G. A. (2012). Across the event horizon. *Current Directions in Psychological Science, 21*(4), 269-272. https://doi.org/10.1177/0963721412451274

Radvansky, G. A., Tamplin, A. K., & Krawietz, S. A. (2010). Walking through doorways causes forgetting: Environmental integration. *Psychonomic Bulletin & Review, 17*(6), 900-904. https://doi.org/10.3758/PBR.17.6.900

Radvansky, G. A., & Zacks, J. M. (2017). Event boundaries in memory and cognition. *Current Opinion in Behavioral Sciences, 17*, 133-140. https://doi.org/10.1016/j.cobeha.2017.08.006

Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D. F., Colombo, J., ... & Thompson, R. F. (2009). Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. *Neurobiology of Learning and Memory*, *92*(2), 135-138. https://doi.org/10.1016/j.nlm.2008.09.012

R Core Team (2021). R: A language and environment for statistical computing (Version 4.0.5) [Computer software]. https://www.R-project.org/

Reed, A. E., & Carstensen, L. L. (2012). The theory behind the age-related positivity effect. *Frontiers in Psychology*, *3*, 30180. https://doi.org/10.3389/fpsyg.2012.00339

Reed, A. E., Chan, L., & Mikels, J. A. (2014). Meta-analysis of the age-related positivity effect: Age differences in preferences for positive over negative information. *Psychology and Aging, 29*(1), 1-15. https://doi.org/10.1037/a0035194

Reichmann, K., Hütter, M., Kaup, B., & Ramscar, M. (2023). Variability and abstraction in evaluative conditioning: Consequences for the generalization of likes and dislikes. *Journal of Experimental Social Psychology*, *108*, 104478. https://doi.org/10.1016/j.jesp.2023.104478

Rescorla, R. A. (1968). Probability of shock in the presence and absence of cs in fear conditioning. *Journal of Comparative and Physiological Psychology, 66*(1), 1–5. https://doi.org/10.1037/h0025984

Rescorla, R. A. (1973). Effects of US habituation following conditioning. *Journal of comparative and physiological psychology*, *82*(1), 137-143. https://doi.org/10.1037/h0033815

Rescorla, R. A. (1974). Effect of inflation of the unconditioned stimulus value following conditioning. *Journal of Comparative and Physiological Psychology*, *86*(1), 101-106. https://doi.org/10.1037/h0035964

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning. Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current Research and Theory.* Appleton Century-Crofts.

Rescorla, R. A. (1991). Associative relations in instrumental learning: The Eighteenth Bartlett Memorial Lecture. *The Quarterly Journal of Experimental Psychology Section B, 43*(1), 1-23.

Revelle, W. (2021). psych: Procedures for Personality and Psychological Research (Version 2.1.6). https://CRAN.R-project.org/package=psych

Richmond, J. L., Zhao, J., & Weidemann, G. (2017). I like it by mere association: Conditioning preferences in infants. *Journal of Experimental Child Psychology*, *161*, 19-31. https://doi.org/10.1016/j.jecp.2017.03.015

Richter, J., & Gast, A. (2017). Distributed practice can boost evaluative conditioning by increasing memory for the stimulus pairs. *Acta Psychologica*, *179*, 1-13. https://doi.org/10.1016/j.actpsy.2017.06.007

Robertson, C. E., Pröllochs, N., Schwarzenegger, K., Pärnamets, P., Van Bavel, J. J., & Feuerriegel, S. (2023). Negativity drives online news consumption. *Nature Human Behaviour, 7*(5), 812-822. https://doi.org/10.1038/s41562-023-01538-4

Ronis, D. L., & Lipinski, E. R. (1985). Value and uncertainty as weighting factors in impression formation. *Journal of Experimental Social Psychology, 21*(1), 47-60. https://doi.org/10.1016/0022-1031(85)90005-8

Rothbart, M., & Park, B. (1986). On the confirmability and disconfirmability of trait concepts. *Journal of Personality and Social Psychology, 50*(1), 131-142. https://doi.org/10.1037/0022-3514.50.1.131

Rothermund, K., & Wentura, D. (2001). Figure-ground asymmetries in the implicit association test (IAT). *Zeitschrift für experimentelle Psychologie*, *48*(2), 94-106. https://doi.orgg/10.1026//0949-3946.48.2.94

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*(2), 225-237. https://doi.org/10.3758/PBR.16.2.225

Rougier, M., & De Houwer, J. (2023). Unconstraining evaluative conditioning research by using the reverse correlation task. *Social Psychological and Personality Science*, 19485506231217526. https://doi.org/10.1177/1948550623121752

Rougier, M., Schmitz, M., & Yzerbyt, V. (2021). When my actions shape your looks: Experience-based properties of approach/avoidance bias the visual representation of others. *Journal of Personality and Social Psychology, 120*(5), 1146-1174. https://doi.org/10.1037/pspa0000268

Rozin, P. (1982). Human food selection: the interaction of biology, culture and individual experience. In L. M. Barker (Ed.), *The psychobiology of human food selection* (pp. 225-254). AVI Publishing Company.

Rozin, P., & Haidt, J. (2013). The domains of disgust and their origins: Contrasting biological and cultural evolutionary accounts. *Trends in Cognitive Sciences, 17*(8), 367-368. https://doi.org/10.1016/j.tics.2013.06.001

Rozin, P., & Millman, L. (1987). Family environment, not heredity, accounts for family resemblances in food preferences and attitudes: A twin study. *Appetite*, *8*(2), 125-134. https://doi.org/10.1016/S0195-6663(87)80005-3

Rozin, P., Millman, L., & Nemeroff, C. (1986). Operation of the laws of sympathetic magic in disgust and other domains. *Journal of Personality and Social Psychology, 50*(4), 703-712. https://doi.org/10.1037/0022-3514.50.4.703

Rozin, P., Wrzesniewski, A., & Byrnes, D. (1998). The elusiveness of evaluative conditioning. *Learning and Motivation*, *29*(4), 397-415. https://doi.org/10.1006/lmot.1998.1012

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*(4), 296-320. https://doi.org/10.1207/S15327957PSPR0504_2

Rubin, E. (1921). Visuell wahrgenommene Figuren (P. Collett, Trans.). Glydendalske. (Original work published 1915)

Ruszpel, B., & Gast, A. (2020). Memory and preparedness in evaluative conditioning in a smell-taste paradigm. A registered report. Cognition and Emotion, 34(5), 1068-1082. https://doi.org/10.1080/02699931.2020.1719982

Rydell, R. J., & Jones, C. R. (2009). Competition between unconditioned stimuli in attitude formation: Negative asymmetry versus spatio-temporal contiguity. *Social Cognition*, *27*(6), 905-916. https://doi.org/10.1521/soco.2009.27.6.905

Sachs, D. H. (1975). Conditioning of affect to a neutral stimulus: Number of trials. *Perceptual and Motor Skills*, *40*(3), 895-901. https://doi.org/10.2466/pms.1975.40.3.8

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language, 110*, Article 104038. https://doi.org/10.1016/j.jml.2019.104038

Schienle, A., Stark, R., & Vaitl, D. (2001). Evaluative conditioning: A possible explanation for the acquisition of disgust responses?. Learning and Motivation, 32(1), 65-83. https://doi.org/10.1006/lmot.2000.1067

Schimmack, U. (2021). The Implicit Association Test: A method in search of a construct. *Perspectives on Psychological Science*, *16*(2), 396-414. https://doi.org/10.1177/1745691619863798

Schmid S., & Poon, C.-S. (2012). *Nonassociative learning.* In N. M. Seel (Ed.), *Encoclypedia of the sciences of learning* (pp. 2475-2477). Springer. http://dx.doi.org/10.1007/978-1-4419-1428-6_1849

Schmidt, S. R. (1991). Can we have a distinctive theory of memory?. *Memory and Cognition*, *19*, 523-542. https://doi.org/10.3758/BF03197149

Schmidt, S. R., & Schmidt, C. R. (2017). Revisiting von Restorff's early isolation effect. *Memory and Cognition, 45*, 194-207. https://doi.org/10.3758/s13421-016-0651-6

Schnabel, K., Asendorpf, J. B., & Greenwald, A. G. (2008). Assessment of individual differences in implicit cognition: A review of IAT measures. *European Journal of Psychological Assessment*, *24*(4), 210-217. https://doi.org/10.1027/1015-5759.24.4.210

Schneider, I. K., & Mattes, A. (2022). The effect of spatial distance between objects on categorization level. *Psychonomic Bulletin & Review, 29*(1), 230-242. https://doi.org/10.3758/s13423-021-01949-2

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*(1-2), 1-46. https://doi.org/10.1016/S0010-0277(00)00152-9

Schwartz, B., Wasserman, E. A., & Robbins, S. J. (2002). *Psychology of learning and behavior*. Norton.

Schwarz, N., & Clore, G. L. (2003). Mood as information: 20 years later. Psychological Inquiry, 14(3-4), 296-303. https://doi.org/10.1080/1047840X.2003.9682896

Seligman, M. E. (1970). On the generality of the laws of learning. *Psychological Review, 77*(5), 406-418. https://doi.org/10.1037/h0029790

Seligman, M. E., & Hager, J. L. (1972). *Biological boundaries of learning.* Appleton-Century-Crofts.

Sherman, J. E. (1978). US inflation with trace and simultaneous fear conditioning. *Animal Learning & Behavior*, *6*(4), 463-468. https://doi.org/10.3758/BF03209645

Sherman, D. K., & Kim, H. S. (2002). Affective perseverance: The resistance of affect to cognitive invalidation. *Personality and Social Psychology Bulletin, 28*(2), 224-237. https://doi.org/10.1177/014616720228200

Shettleworth, S. J. (2009). *Cognition, evolution, and behavior*. Oxford University Press.

Shimp, T. A., Stuart, E. W., & Engle, R. W. (1991). A program of classical conditioning experiments testing variations in the conditioned stimulus and context. *Journal of Consumer Research*, *18*(1), 1-12. https://doi.org/10.1086/209236

Shin, Y. S., & DuBrow, S. (2021). Structuring memory through inference-based event segmentation. *Topics in Cognitive Science, 13*(1), 106-127. https://doi.org/10.1111/ tops.12505

Simonsohn, U. (2014). [17] No-way interactions. https://doi.org/10.15200/ winn.142559.90552

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2015). Better P-curves: Making P-curve analysis more robust to errors, fraud, and ambitious P-hacking, a Reply to Ulrich and Miller (2015). *Journal of Experimental Psychology: General, 144*(6), 1146-1152. https://doi.org/10.1037/xge0000104

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Sachar, M. S. (2021). afex: Analysis of Factorial Experiments (Version 1.0-1). https://CRAN.R-project.org/package=afex

Skinner, B. F. (2019). *The behavior of organisms: An experimental analysis*. BF Skinner Foundation. (Original work published 1938)

Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology, 52*(4), 689-699. https://doi.org/10.1037/0022-3514.52.4.689

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. Psychological Bulletin, 105(1), 131-142. https://doi.org/10.1037/0033-2909.105.1.131

Sparks, J., & Ledgerwood, A. (2019). Age attenuates the negativity bias in reframing effects. *Personality and Social Psychology Bulletin*, *45*(7), 1042-1056. https://doi.org/10.1177/0146167218804526

Sperlich, L. M., & Unkelbach, C. (2022). When do people learn likes and dislikes from co-occurrences? A dual-force perspective on evaluative conditioning. *Journal of Experimental Social Psychology*, *103*, 104377. https://doi.org/10.1016/j.jesp.2022.104377

Staats, A. W., & Staats, C. K. (1958). Attitudes established by classical conditioning. *The Journal of Abnormal and Social Psychology, 57*(1), 37-40. https://doi.org/10.1037/h0042782

Staats, A. W., & Staats, C. K. (1959). Effect of number of trials on the language conditioning of meaning. *The Journal of General Psychology*, *61*(2), 211-223. https://doi.org/10.1080/00221309.1959.9710255

Staats, A. W., Staats, C. K., & Heard, W. G. (1959). Language conditioning of meaning using a semantic generalization paradigm. *Journal of Experimental Psychology, 57*(3), 187-192. https://doi.org/10.1037/h0042274

Stahl, C., & Aust, F. (2018). Evaluative conditioning as memory-based judgment. *Social Psychological Bulletin*, *13*(3), 1-30. https://doi.org/10.5964/spb.v13i3.28589

Stahl, C., & Bading, K. C. (2020). Evaluative conditioning of pattern-masked nonwords requires perceptual awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(5), 822-850. https://doi.org/10.1037/xlm0000757

Stahl, C., Haaf, J., & Corneille, O. (2016). Subliminal evaluative conditioning? Above-chance CS identification may be necessary and insufficient for attitude learning. *Journal of Experimental Psychology: General, 145*(9), 1107-1131. https://doi.org/10.1037/xge0000191

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

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

Stieger, S., & Burger, C. (2013). More complex than previously thought: New insights into the optimal administration of the initial preference task. *Self and Identity*, *12*(2), 201-216. https://doi.org/10.1080/15298868.2012.655897

Stieger, S., Voracek, M., & Formann, A. K. (2012). How to administer the initial preference task. *European Journal of Personality*, *26*(1), 63-78. https://doi.org/10.1002/per.823

Stuart, E. W., Shimp, T. A., & Engle, R. W. (1987). Classical conditioning of consumer attitudes: Four experiments in an advertising context. *Journal of Consumer Research*, *14*(3), 334-349. https://doi.org/10.1086/209117

Suls, J., Martin, R., & Wheeler, L. (2002). Social comparison: Why, with whom, and with what effect? *Current Directions in Psychological Science, 11*(5), 159-163. https://doi. org/10.1111/1467-8721.00191

Sweatt, J. D. (2009). *Mechanisms of memory*. Academic Press.

Sweldens S. (2018). *Putting operating principles (what) before conditions (how and when) to improve theorizing: S–S versus S–R learning in evaluative conditioning*. Retrieved from osf.io/bu4v5

Sweldens, S., Corneille, O., & Yzerbyt, V. (2014). The role of awareness in attitude formation through evaluative conditioning. *Personality and Social Psychology Review*, *18*(2), 187-209. https://doi.org/10.1177/1088868314527832

Sweldens, S., Van Osselaer, S. M., & Janiszewski, C. (2010). Evaluative conditioning procedures and the resilience of conditioned brand attitudes. *Journal of Consumer Research*, *37*(3), 473-489. https://doi.org/10.1086/653656

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilizationminimization hypothesis. *Psychological Bulletin, 110*(1), 67-85. https://doi.org/10.1037/0033- 2909.110.1.67

Tello, N., Bocage-Barthélémy, Y., Dandaba, M., Jaafari, N., & Chatard, A. (2018). Evaluative conditioning: A brief computer-delivered intervention to reduce college student drinking. *Addictive Behaviors, 82,* 14-18. https://doi.org/10.1016/j.addbeh.2018.02.018

Tesser, A. (1978). Self-generated attitude change. *Advances in Experimental Social Psychology, 11*, 289-338. https://doi.org/10.1016/S0065-2601(08)60010-6

Theeuwes, J. (2018). Visual selection: Usually fast and automatic; seldom slow and volitional. Journal of Cognition, 1, 1–15. https://doi.org/10.5334/joc.13 Unkelbach, C., Alves, H., & Koch, A. (2020). Negativity bias, positivity bias, and valence asymmetries: Explaining the differential processing of positive and negative asymmetries: Explaining the differential processing of positive and negative information. *Advances in Experimental Social Psychology, 62*, 115-187. https://doi. org/10.1016/bs.aesp.2020.04.005

Thompson, R. F. (2009). Habituation: a history. *Neurobiology of learning and memory*, *92*(2), 127-134. https://doi.org/10.1016/j.nlm.2008.07.011

Thompson, R. F., & Glanzman, D. L. (2016). Neural and behavioral mechanisms of habituation and sensitization. In T. J. Tighe, & R. N. Leaton, *Habituation* (pp. 49-93). Routledge. https://doi.org/10.4324/9781315636726

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *The Psychological Review: Monograph Supplements, 2*(4), 1-109. https://doi.org/10.1037/h0092987

Thorpe, W. H. (1956). *Learning and instinct in animals.* Harvard University Press.

Till, B. D., & Priluck, R. L. (2000). Stimulus generalization in classical conditioning: An initial investigation. Psychology & Marketing, 17, 55-72. https://doi.org/10.1002/(SICI)1520-6793(200001)17:1<55::AID-MAR4>3.0.CO;2-C

Todrank, J., Byrnes, D., Wrzesniewski, A., & Rozin, P. (1995). Odors can change preferences for people in photographs: A cross-modal evaluative conditioning study with olfactory USs and visual CSs. *Learning and Motivation*, *26*(2), 116-140. https://doi.org/10.1016/0023-9690(95)90001-2

Tolman, E. C., & Honzik, C. H. (1930). "Insight" in rats. *University of California Publications in Psychology, 4*, 215-232.

Topolinski, S., & Deutsch, R. (2013). Phasic affective modulation of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(2), 414-436. https://doi.org/10.1037/a0028879

Torchiano, M. (2020). effsize: Efficient Effect Size Computation (Version 0.8.1). https://CRAN.R-project.org/package=effsize

Treisman, A. (1986). Features and objects in visual processing. *Scientific American, 255*(5), 114B-125. https://www.jstor.org/stable/24976089

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97-136. https://doi.org/10.1016/0010-0285(80)90005-5

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics, 106*(4), 1039-1061. https://doi.org/10.2307/2937956

Unkelbach, C., Alves, H., & Koch, A. (2020). Negativity bias, positivity bias, and valence asymmetries: Explaining the differential processing of positive and negative information. *Advances in Experimental Social Psychology, 62*, 115-187. https://doi.org/10.1016/bs.aesp.2020.04.005

Unkelbach, C., & Fiedler, K. (2016). Contrastive CS-US relations reverse evaluative conditioning effects. *Social Cognition*, *34*(5), 413-434. https://doi.org/10.1521/soco.2016.34.5.413

Unkelbach, C., Fiedler, K., Bayer, M., Stegmüller, M., & Danner, D. (2008). Why positive information is processed faster: The density hypothesis. *Journal of Personality and Social Psychology, 95*(1), 36-49. https://doi.org/10.1037/0022-3514.95.1.36

Unkelbach, C., & Högden, F. (2019). Why does George Clooney make coffee sexy? The case for attribute conditioning. *Current Directions in Psychological Science*, *28*(6), 540-546. https://doi.org/10.1177/0963721419859354

Unkelbach, C., Koch, A., & Alves, H. (2019). The evaluative information ecology: On the frequency and diversity of "good" and "bad". *European Review of Social Psychology*, *30*(1), 216-270. https://doi.org/10.1080/10463283.2019.1688474

Unkelbach, C., Koch, A., & Alves, H. (2021). Explaining negativity dominance without processing bias. *Trends in Cognitive Sciences, 25*(6), 429-430. https://doi.org/10.1016/j.tics.2021.04.005

Unkelbach, C., Ostheimer, V., Fasold, F., & Memmert, D. (2012). A calibration explanation of serial position effects in evaluative judgments. *Organizational Behavior and Human Decision Processes, 119*(1), 103-113. https://doi.org/10.1016/j. obhdp.2012.06.004

Unkelbach, C., Stahl, C., & Förderer, S. (2012). Changing CS features alters evaluative responses in evaluative conditioning. *Learning and Motivation, 43*(3), 127-134. https://doi.org/10.1016/j.lmot.2012.04.003

Van Dessel, P., Cummins, J., Hughes, S., Kasran, S., Cathelyn, F., & Moran, T. (2020). Reflecting on 25 years of research using implicit measures: Recommendations for their future use. *Social Cognition*, *38*(Supplement), 223-242. https://doi.org/10.1521/soco.2020.38.supp.s223

Van Dessel, P., Hughes, S., & De Houwer, J. (2019). How do actions influence attitudes? An inferential account of the impact of action performance on stimulus evaluation. *Personality and Social Psychology Review*, *23*(3), 267-284. https://doi.org/10.1177/10888683187957

Vansteenwegen, D., Francken, G., Vervliet, B., De Clercq, A., & Eelen, P. (2006). Resistance to extinction in evaluative conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 32*(1), 71-79. https://doi.org/10.1037/0097-7403.32.1.71

Veltkamp, M., Custers, R., & Aarts, H. (2011). Motivating consumer behavior by subliminal conditioning in the absence of basic needs: Striking even while the iron is cold. *Journal of Consumer Psychology*, *21*(1), 49-56. https://doi.org/10.1016/j.jcps.2010.09.011

Verwijmeren, T., Karremans, J. C., Stroebe, W., & Wigboldus, D. H. (2012). Goal relevance moderates evaluative conditioning effects. *Learning and Motivation*, *43*(3), 107-115. https://doi.org/10.1016/j.lmot.2012.06.002

Vianello, M., & Bar-Anan, Y. (2021). Can the Implicit Association Test measure automatic judgment? The validation continues. *Perspectives on Psychological Science*, *16*(2), 415-421. https://doi.org/10.1177/1745691619897960

Vogel, T., Hütter, M., & Gebauer, J. E. (2019). Is evaluative conditioning moderated by Big Five personality traits?. *Social Psychological and Personality Science*, *10*(1), 94-102. https://doi.org/10.1177/1948550617740193

Vogel, T., & Wänke, M. (2016). Attitudes and attitude change. Psychology Press. https://doi.org/10.4324/9781315754185

Von Restorff, H. (1933). Über die Wirkung von Bereichsbildungen im Spurenfeld. *Psychologische Forschung, 18*, 299-342. https://doi.org/10.1007/BF02409636

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin, 138*(6), 1172-1217. https://doi.org/10.1037/a0029333

Walasek, L., & Stewart, N. (2015). How to make loss aversion disappear and reverse: Tests of the decision by sampling origin of loss aversion. Journal of Experimental Psychology: General, 144(1), 7-11. https://doi.org/10.1037/xge0000039

Walther, E. (2002). Guilty by mere association: Evaluative conditioning and the spreading attitude effect. *Journal of Personality and Social Psychology, 82*(6), 919-934. https://doi.org/10.1037/0022-3514.82.6.919

Walther, E., Halbweisen, G., & Blask, K. (2018). What you feel is what you see: A binding perspective on evaluative conditioning. *Social Psychological Bulletin*, *13*(3), 1-18. https://doi.org/10.5964/spb.v13i3.27551

Walther, E., & Grigoriadis, S. (2004). Why sad people like shoes better: The influence of mood on the evaluative conditioning of consumer attitudes. *Psychology & Marketing*, *21*(10), 755-773. https://doi.org/10.1002/mar.20028

Walther, E., Langer, T., Weil, R., & Komischke, M. (2011). Preferences surf on the currents of words: Implicit verb causality influences evaluative conditioning. *European Journal of Social Psychology*, *41*(1), 17-22. https://doi.org/10.1002/ejsp.785

Walther, E., & Langer, T. (2008). Attitude formation and change through association: An evaluative conditioning account. In W. D. Crano, & R. Prislin (Eds.), *Attitudes and attitude change* (pp. 87-109). Taylor & Francis Group.

Walther, E., & Langer, T. (2010). For whom Pavlov's bell tolls: Processes underlying evaluative conditioning. In J. P. Forgas, J. Cooper, & W. D. Crano (Eds.), *The psychology of attitudes and attitude change* (pp. 59-74). Taylor & Francis Group.

Walther, E., & Nagengast, B. (2006). Evaluative conditioning and the awareness issue: Assessing contingency awareness with the Four-Picture Recognition Test. *Journal of Experimental Psychology: Animal Behavior Processes, 32*(4), 454-459. https://doi.org/10.1037/0097-7403.32.4.454

Walther, E., Nagengast, B., & Trasselli, C. (2005). Evaluative conditioning in social psychology: Facts and speculations. *Cognition and Emotion*, *19*(2), 175-196. https://doi.org/10.1080/02699930441000274

Wardle, S. G., Mitchell, C. J., & Lovibond, P. F. (2007). Flavor evaluative conditioning and contingency awareness. *Learning & Behavior*, *35*, 233-241. https://doi.org/10.3758/BF03206429

Warriner, A. B., & Kuperman, V. (2015). Affective biases in English are bi-dimensional. *Cognition and Emotion*, *29*(7), 1147-1167. https://doi.org/10.1080/02699931.2014.968098

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92*(4), 548-573. https://doi.org/10.1037/0033-295X.92.4.548

Weiner, B., & Graham, S. (1984). An attributional approach to emotional development. In C. E. Izard, J. Kagan, & R. B. Zajonc (Eds.), *Emotions, cognition, and behavior* (pp. 167-191). Cambridge University Press.

Wertheimer, M. (1938). Laws of organization in perceptual forms (partial translation). In W. D. Ellis (Ed.), *A source book of gestalt psychology* (pp. 71-88). Hartcourt Brace. https://doi.org/10.1037/11496-005

Wickham, H. (2007). Reshaping Data with the reshape Package. Journal of Statistical Software, 21(12), 1-20. https://10.18637/jss.v021.i1

Wickham, H. (2011). The Split-Apply-Combine Strategy for Data Analysis. Journal of Statistical Software, 40(1), 1-29. https://10.18637/jss.v040.i01

Wickham, H. (2019). stringr: Simple, Consistent Wrappers for Common String Operations (Version 1.4.0). https://CRAN.R-project.org/package=stringr

Wickham, H., Henry, L., Vaughan, D. (2023). vctrs: Vector Helpers (Version 0.6.5). https://github.com/r-lib/vctrs

Wickham, H., François, R., Henry, L., & Müller, K. (2021). dplyr: A Grammar of Data Manipulation (Version 1.0.7). https://CRAN.R-project.org/package=dplyr

Wickham, H. & Hester, J. (2021). readr: Read Rectangular Text Data (Version 2.0.1). https://CRAN.R-project.org/package=readr

Wickham H., Vaughan D., Girlich M. (2024). tidyr: Tidy Messy Data (Version 1.3.1). https://github.com/tidyverse/tidyr

White, K., & Davey, G. C. (1989). Sensory preconditioning and UCS inflation in human 'fear'conditioning. *Behaviour Research and Therapy, 27*(2), 161-166. https://doi.org/10.1016/0005-7967(89)90074-0

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*(2, Pt.2), 1-27. https://doi.org/10.1037/h0025848

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

Zellner, D. A., Rozin, P., Aron, M., & Kulish, C. (1983). Conditioned enhancement of human's liking for flavor by pairing with sweetness. *Learning and Motivation*, *14*(3), 338-350. https://doi.org/10.1016/0023-9690(83)90021-8

Zerhouni, O., Bègue, L., Comiran, F., & Wiers, R. W. (2018). Controlled and implicit processes in evaluative conditioning on implicit and explicit attitudes toward alcohol and intentions to drink. *Addictive behaviors*, *76*, 335-342. https://doi.org/10.1016/j.addbeh.2017.08.026

Zerhouni, O., & Lepage, J. (2018). "Paint it black": Chronic emotion regulation difficulties enhances evaluative conditioning effect when conditioned stimuli are paired with highly arousing negative unconditioned stimuli. *Journal of Individual Differences, 39*(4), 196-203. https://doi.org/10.1027/1614-0001/a000264

Zorn, T. J., Mata, A., & Alves, H. (2022). Attitude similarity and interpersonal liking: A dominance of positive over negative attitudes. *Journal of Experimental Social Psychology*, *100*, 104281. https://doi.org/10.1016/j.jesp.2021.104281