How associative is Evaluative Conditioning?

An empirical investigation at the attitude acquisition and attitude change stages

How associative is Evaluative Conditioning? An empirical investigation at the attitude acquisition and attitude change stages

Adrien Mierop

Promoteur

Olivier Corneille (UCL)

Thèse présentée en vue de l'obtention du grade de Docteur en sciences psychologiques et de l'éducation

Comité d'encadrement

Christoph Stahl (Universität zu Köln)
Nicolas Vermeulen (UCL)

UCL PRESSES UNIVERSITAIRES DE LOUVAIN

Collection de thèses de l'Université catholique de Louvain, 2018

Président

Pierre Maurage

Jury :

O. Corneille, J. De Houwer, M. Hütter, C. Stahl, N. Vermeulen

Table of Contents

Abstract

This doctoral thesis examines the contribution of associative learning to evaluative conditioning effects (EC). Associative learning in attitude research is assumed to involve the automatic formation of unqualified links between conditioned (CS) and unconditioned (US) stimuli in memory. We addressed the « automaticity » assumption by examining the role of awareness in EC, and the efficiency of EC. The « unqualified » assumption was addressed by examining the reciprocity of US-to-CS and CS-to-US effects following CS-US pairings, and the sensitivity of EC to a change in CS-US regularities. Results from multiple experiments question both the automatic and unqualified EC assumptions. Therefore, this doctoral research challenges the view that associative learning processes play a role in EC and, more generally, the validity of dual-learning models of attitudes.

Résumé

Cette thèse de doctorat porte sur la contribution de l’apprentissage associatif aux effets de conditionnement évaluatif (CE). L'apprentissage associatif dans la recherche sur les attitudes est supposé impliquer la formation automatique de liens non qualifiés entre les stimuli conditionnés (SC) et inconditionnés (SI) en mémoire. Nous avons abordé l'hypothèse d'« automaticité » en examinant le rôle de la conscience dans le CE et l'efficience du CE. L’hypothèse de la « non qualification » a été examinée en examinant la réciprocité des effets entre SC et SI lors de leurs appariements et la sensibilité du CE à une modification des régularités entre SC et SI. Les résultats de plusieurs expériences remettent en question à la fois les hypothèses d’automaticité et de non-qualification. Par conséquent, cette recherche doctorale remet en cause l'idée selon laquelle des processus d'apprentissage associatifs jouent un rôle dans le CE et, plus généralement, la validité des modèles duaux de formation d’attitude.

Acknowledgments/Remerciements

J'ai passé quatre ans sur ma thèse, quatre ans de découvertes scientifiques mais aussi humaines. Quand je regarde en arrière je remarque à quel point cette expérience a été intense et m'a changé. Je tiens ici à remercier les personnes qui y ont participé. Je pourrais écrire autant de pages de remerciements que de thèse mais je tenterai de rester succinct.

En premier lieu je tiens à remercier Olivier. Il est clairement conscient que je lui suis énormément reconnaissant, je ne lui dirai jamais assez… Je suis vraiment content d'avoir l'occasion de développer un peu ici.

J'ai rencontré Olivier à son cours sur le changement d'attitude. J'étais alors un étudiant moyen voire carrément médiocre qui ne savait pas trop ce qu'il faisait là. J'étais loin de me douter que 7 ans plus tard je rendrais une thèse sous sa supervision précisément sur le changement d'attitude. Olivier a été (et est toujours) une personne clé dans ma vie. Et sincèrement, je pèse mes mots. Il m'a donné une passion et m'a aidé à pouvoir en vivre. Au-delà de ça, il est une très belle personne dont je suis honoré de partager un bout de sa vie.

J'ai commencé à travailler avec Olivier par un stage en recherche de cinq mois. Olivier a tout de suite placé la barre très haut. Après l'avoir rencontré pour la première fois, je suis ressorti avec un énorme ouvrage « implicit social cognition » (le genre de livre absolument indigeste et d'une densité phénoménale pour un étudiant fraichement débarqué de Bachelier) et une (très forte) recommandation à avoir « au strict minimum » une grande distinction si je veux continuer en recherche… Moi qui sortait de mon bac en ayant doublé… Lors de ce stage, il a transformé mon intérêt pour la recherche en une passion. Il m'a alors proposé de faire un mémoire avec lui. Pour ce mémoire, il m'a laissé une liberté incroyable tout en faisant preuve d'une exigence toujours plus grande : « Olivier, l'exigence de l'excellence ». Olivier a vu que j'étais sérieux dans ma motivation de faire de la recherche et m'a alors soutenu pour déposer un projet de thèse avant même d'avoir terminé mon master. Olivier avait les idées, l'expérience, la clairvoyance, le réseau, la réputation. Le projet est passé « comme une lettre à la poste ».

Dès le début de ma thèse, il m'a appris à réfléchir et à poser de bonnes questions. Je me souviens au tout début, j'arrivais souvent dans son bureau « Olivier, j'ai 16 idées dont j'aimerais parler avec toi ». Je ressortais rarement avec ne fut-ce qu'une idée avec laquelle on pouvait avancer. Mais il a toujours pris le temps de les écouter, de les commenter ou les raffiner.

J'ai beaucoup de chance d'avoir Olivier comme promoteur, il a été un encadrant idéal pour moi. D'un côté, il garde toujours le « big picture » en vue et me redirige toujours sur les projets qui sont *vraiment* important. D'un autre, il m'a donné beaucoup d'autonomie et m'a fait confiance sur une série de projets qui finalement servaient surtout à assouvir ma curiosité. C'était très important pour moi car par cet équilibre, j'apprenais à faire de la recherche et je prenais du plaisir à ça.

Il a aussi un équilibre qui me correspond vraiment bien vis-à-vis des feedbacks qu'il me donne. Quand je disais qu'Olivier est très exigeant, ce n'est pas une blague. Il l'est énormément. Pas une fois en écrivant un article il ne m'a pas mis face à mon insuffisance. Pas du tout méchamment, au contraire, c'est une marque de respect et de considération. Ma manière de communiquer, de réfléchir, les designs expérimentaux, ..., tout devait toujours être mieux. Je ne vois pas un seul moment dans ma thèse durant lequel j'ai stagné en termes de compétence, j'ai toujours été forcé de m'améliorer. D'un autre côté, Olivier m'a très souvent félicité pour ce qu'il estimait être bon (pour lui quand quelque chose est « bon », c'est que c'est très très bon...). Ça allait d'une simple réponse à un mail, à un paragraphe écrit dans tel ou tel article, à une idée en discussion de groupe, etc. Il m'a aussi plusieurs fois dit qu'il était très content de travailler avec moi, tant pour le côté de la recherche que pour le côté humain. De nouveau, cet équilibre me permettait de progresser en prenant du plaisir à me dépasser.

Olivier m'a aussi très vite incité à participer à des conférences et m'a présenté à de nombreuses personnes qui, elles aussi, seront clés pour moi. Dès le début de ma thèse, j'ai commencé à travailler avec Mandy, Christoph et Nicolas que je remercierai dans un instant.

Au-delà de ses qualités en tant que chercheur et promoteur, Olivier a de très belles qualités humaines. Olivier est quelqu'un de sincèrement bienveillant. C'est super rare, mais il se soucie *vraiment* du bien-être des autres. Attention, il peut être dur dans le sens où il ne va pas toujours dire des choses qu'on veut entendre, mais toujours dans l'intérêt de l'autre. A plusieurs moments durant ma thèse j'ai eu des moments où j'allais moins bien. Il a très vite reconnu que ça n'allait pas et, systématiquement, prenait le temps de discuter avec moi pour comprendre et me donner sa perspective. J'ai toujours été très surpris de sa maturité dans les relations interpersonnelles.

Je repense aussi à ce tout ce qu'Olivier a fait pour moi et c'est assez scandaleux à quel point j'ai de la chance en fait… Olivier m'a permis d'avoir une bourse de thèse directement après mon master et là, il vient de décrocher une bourse sur laquelle je peux encore travailler en post-doctorat pendant deux ans… Olivier est déjà en train de penser à un autre post-doctorat après celui-ci et a contacté la doyenne pour savoir comment je peux faire pour donner des cours (je dois rassembler un certain nombre d'heures de cours pour pouvoir postuler à des postes).

Pour toutes ces choses et bien plus, je tiens à te remercier Olivier, j'ai beaucoup de chance et j'espère pouvoir continuer à travailler (et manger des pizzas) avec toi le plus longtemps possible.

Je tiens également à remercier très chaleureusement les membres de mon comité d'accompagnement : Nicolas Vermeulen et Christoph Stahl. Nicolas particulièrement pour son enthousiasme par rapport à la recherche et sa créativité. C'était un réel plaisir de discuter avec lui sur un paradigme en particulier (attentional blink) et comment l'adapter au conditionnement. Nicolas a l'œil qui s'allume quand il parle de recherche. Pour mon mémoire, je suis allé discuter avec lui pour développer une mesure comportementale. Il a trouvé quasi instantanément trois idées différentes et est resté discuté avec moi longuement sur mon hypothèse théorique. Sentir ce plaisir pour la recherche m'avait vraiment inspiré.

I would also like to thank Christoph for all the discussions and collaborations we had. A really interesting thing with Christoph is how he manages to provide a comment, a criticism or a question that leads you to rethink deeply about your work. It usually took me some time to actually understand and integrate the comments. For instance, during my private defence, he asked me to specify my level of analysis. I was a bit confused and was not really able to provide an answer at that point. It stayed in my mind for a while and I finally understood (or at least I believe I understood) what he was asking. I cannot think without distinguishing the levels of analysis by now. I feel truly honored to get to work with him.

I also wish to thank the (other) members of the jury of my thesis : Mandy Hütter and Jan De Houwer. I have of course a lot to say about Mandy, I collaborated with her for almost the entire duration of my thesis and I spent six months working with her at her lab in Tübingen. Mandy is the perfect example that someone can be both competent and warm. She is very impressive in her work. I hardly saw someone as organized as she is. I sometimes have the impression that she has a twin and even though the two Mandys would be very impressive. Mandy is very exigent and very constructive in her comments. She never points to a problem without suggesting a solution. Mandy is also someone you can easily talk to, being for fun or for more personal reasons. Mandy is also the type of person that will call you for one hour for your birthday discussing about everything. I felt very welcome in Tübingen and for that as well I want to thank her. I sincerely hope I will be able to continue working with her for long.

I would like to thank Jan as well. Jan is clearly a pioneer researcher in his domain. I feel honered having received his feedbacks on my dissertation as well as for several articles. Jan is for me an example of humility and constructivism in research.

Je tiens également à remercier Pierre Maurage pour avoir assuré la présidence de mon jury de thèse. Ses commentaires étaient très stimulants et j'espère sincèrement avoir l'occasion de travailler avec lui lors de ces prochaines années.

Je souhaite maintenant remercier tout.e.s les collègues qui m'ont accompagné.e.s durant cette thèse (dans un ordre plus ou moins chronologique).

Je tiens à remercier chaleureusement Vincent Yzerbyt (pour toi aussi démontrer que oui, on peut être compétent ET chaleureux, pour ton implication dans la recherche, dans la communauté scientifique, et ton « need for cognition » dont j'ai pu attester à plusieurs reprises que ce soit en réunion de laboratoire, en voiture au retour de conférence ou encore derrière un script R, et pour m'avoir donné le goût pour les statistiques), Florence Stinglhamber (pour ton humour, ton esprit de l'organisation et m'avoir convaincu Saint-Nicolas existe bel et bien !), Stéphanie Demoulin (pour ta qualité d'enseignement et plus généralement de communication, pour avoir partagé de très chouettes moments et m'avoir convaincu que partager une orange n'implique pas nécessairement de privation), Bernard Rimé (pour m'avoir très tôt encouragé, pour représenter ce qu'un chercheur peut fonder en une carrière académique, pour avoir eu plaisir à t'écouter quand tu présentes tes travaux avec passion), Ginette Herman (pour ton éternelle bonne humeur, ton éternel appétit aussi…, et pour ton sincère soucis des autres), Donatienne Desmette (ta bonne humeur contagieuse et ton calme indéfectible), Vassilis Saroglou (pour votre profond intérêt pour la recherche, pour m'avoir secoué en baccalauréat et m'avoir donné un aperçu de la psychologie en tant que discipline scientifique), Rafaele Dumas (pour nos nombreux débats et nos nombreux verres), et Annalisa Casini (pour ton franc parler, ton fort sens de l'humour et tes convictions).

Je tiens aussi à remercier Catia (aka « Cookie ») qui a partagé non seulement son amour pour la recherche mais aussi presque deux ans avec moi. C'est évidemment difficile de te remercier ici en seulement quelques lignes. Je m'estime très chanceux d'avoir pu partager tout ce temps avec toi. J'ai beaucoup appris pendant cette période, j'ai rarement été aussi stimulé sur le plan de la recherche et j'ai, accessoirement…, appris à me connaître. On a passé plein de super bons moments qui m'ont clairement fait devenir ce que je suis aujourd'hui. On a découvert un continent ensemble, j'ai appris à faire un site pour pouvoir te convaincre, j'ai testé tous les *frango no churrasco* de Bruxelles et de Porto. J'ai eu tellement de fous rires avec toi, j'en rigole en l'écrivant. Merci beaucoup pour tout, je suis fier de pouvoir te compter comme une de mes vraies amies.

Je tiens aussi à remercier Timothée et Gaylord (pour m'avoir encadré lors de mon stage de recherche qui a clairement déterminé mon parcours, j'ai eu beaucoup de chance de tomber sur vous deux), les plus anciens comme Jonathan, Edwine, Evelyne, Caro, Marie, Séverine, Claire (pour m'avoir accueilli chaleureusement dans votre couloir), ainsi que toutes les personnes du troisième avec qui j'ai passé de merveilleuses années. Hélène (pour ton caractère, pour ton amour de la Belgique et tout ce qui va avec), Tina (pour ta bonne humeur, tes convictions), Nolwenn (pour ton ensoleillement, pour tous nos délires et nos verres), Mathias S. (pour ton attitude « all-in », ta passion pour la programmation et ton besoin d'absolument TOUT comprendre), Neptina (pour ton sourire et ta bienveillance), Nathan (pour ton rire permanent, pour ton amour du sport).

Je tiens aussi à remercier mes docteur.e.s français préféré.e.s, Antonin (pour tous nos moments à discuter de philo, de musique, de recherche, de voyages), Marine (ou plutôt « docteure Rougier », pour ton amour des traces, des stats, pour ton humour… piquant mais tellement bon), Jen (pour ton altruisme, pour notre alcoolémie identique, pour cette merveilleuse « révélation » de l'année).

Je tiens vraiment à remercier Julie et Amélie. Julie pour ton humeur si harmonieuse… Pour ta sincérité, tout ce que tu fais pour les autres. Pour ton caractère unique. Pour ton talent et ta proactivité. Amé c'est difficile d'écrire ici toutes les choses pour lesquelles je te suis reconnaissant. Tu m'as ouvert. Ouvert à la programmation, à l'open-science mais aussi à l'art, à la « cohérence » dans le mode de vie quotidien, au moment présent. Tu m'as aussi ouvert à moi-même. J'ai conscience avoir fondamentalement changé grâce à toi et je t'en remercie vraiment… Je n'oublie pas Amé.

Je dois évidemment remercier les trois autres membres des « quatre fantastiques » : Gaëtane, Matthias D. et Christophe. Gaet particulièrement pour ton franc-parler, tu ne passes pas par quatre chemins quand tu veux dire quelque chose. J'ai découvert que derrière cette Gaëtane dure et critique se cache une très belle personne. On s'est tous les deux vus dans des bons et des mauvais moments, j'ai toujours eu l'impression qu'on se comprenait sans nécessairement avoir besoin de trop parler. Tu sais ce que tu veux et tu y vas, c'est quelque chose que je respecte beaucoup. Mat pour le fait que tu repenses constamment les choses les plus basiques. Pour tous nos fous rires, particulièrement dans les moments où on ne peut pas de tout se le permettre… Pour notre intérêt commun très gamin pour les jeux-vidéos et pour ton attitude tout-ou-rien en soirée. On a toujours pu se dire les choses directement et on est tous les deux sincèrement intéressé par l'autre. Evidemment Chris… « je vais peut-être pleurer » en te remerciant. Je te remercie vraiment pour tout. Je ne sais pas comment se serait passé ma thèse sans toi en fait. J'ai vraiment découvert quelqu'un de super et je suis fier d'être devenu ton ami. Merci pour avoir été là dans les meilleurs moments et dans les pires « mec… j'ai déconné ». Je te remercie aussi pour tous les moments passés en mer avec Devil Peavey à guetter l'approche du kraken. Pour avoir créé l'Umbra Consortium. Pour nos soirées qui dégénèrent en route du rhum. Pour nos cafés le matin et pour notre humour commun pour les trucs débiles. Rien que pour cette vraie amitié, ces quatre années de thèse en valaient le coup.

I also wish to thank the members of the lab in Tübingen who welcomed me for six months. Max, Kathi, Fabian, Maximilian, and Neils. I was very lucky to meet you and learn from you. I was also lucky to discover new things such as a blind test of spezi. My only regret is that I couldn’t make the Alpaca hike with you guys… I wish to especially thank Maximilian who became a true friend. I rarely had so rich discussions, both for serious and less-serious matters. I learnt a lot by discussing with you, I also discovered meditation that changed my life.

J’aimerais aussi remercier tous les gens qui m’ont accompagné durant ces quatre années dans une sphère moins professionnelle. Premièrement les « anciens », Ollie, Maxime, Quentin, Ludo, Forence, Morgane 1, Morgane 2, Delphine, Julien, Salva, Karim et Thibaud. On a grandi ensemble et vous m’avez clairement dirigé vers la psychologie et la recherche. Également les « chefs loups » Vincent, Blaise, Lola, Matthieu et Félix avec qui j’ai passé beaucoup de temps au moment où je commençais à réellement vouloir étudier. On a passé plein de super moments qui resteront gravés dans ma mémoire. J’aimerais aussi remercier les « musicos » Holybass, Gilles, Albert, Quentin et Sammy avec qui j’ai passé des supers soirées et ai pu m’amuser avec une autre passion : les percussions. Enfin, je tiens à remercier Sophie avec qui j’ai passé quatre mois magnifiques et avec qui j’ai découvert l’Equateur. C’étaient de très beaux moments et pour ça je te remercie.

Je tiens évidemment à remercier chaleureusement ma famille qui m’a toujours soutenu durant mes études et dans mon projet de faire une thèse. Je devrais vous remercier tous et toutes ici mais la liste est vraiment trop longue, donc je vous remercierai plutôt à l’oral. Je tiens tout de même à dire quelques mots pour ma « petite » famille. D’abord ma mère qui m’a toujours fait garder la tête froide dans des moments tendus, qui a l’art de toujours me faire relativiser de ce qui se passe dans ma vie et qui se soucie toujours de mon moral plutôt que de mes avancées. Ensuite mon père qui m’a toujours super motivé dans la recherche. Il est toujours désireux de savoir comment se passe tel ou tel projet, où en est cet article « qui doit révolutionner le monde » et de parler de R et des stats en dégustant un Amarone. Je tiens tout particulièrement remercier ma petite sœur adorée qui a vécu ma thèse de presque aussi près que moi. Elle aussi m’a soutenu dans des moments assez difficiles, j’ai beaucoup de chance que tu partages ma vie ! Je te remercie aussi pour tous nos délires et notre humour pourri, mais vraiment pourri…, nos parties de jeux vidéo avec bières et pizzas. Rock and Stone Sister !

J'ai énormément de chance de vous avoir, et je vous remercie du fond du cœur pour tout ce que vous êtes pour moi.

General introduction

Chapter One: General introduction

Evaluative Conditioning can be defined as a change in valence of a stimulus (i.e., the conditioned stimulus, or CS) resulting from its pairing with an affective stimulus (i.e., the unconditioned stimulus, or US) (De Houwer, Thomas, & Baeyens, 2001). It is important to distinguish the procedure (i.e., the CS-US pairing), the effect (i.e., the change in valence) and the processes through which an evaluative conditioning procedure results in an evaluative effect. Unless otherwise stated, I will consider here EC as an effect (De Houwer, 2007). This choice allows considering that EC is determined by multiple processes. The experiments presented in this dissertation will be all addressing the contribution of an associative learning process to EC.

In the introduction, I first describe EC and the origin of its investigation. I then sketch its applied potential in a variety of domains, including marketing, politics, health psychology and intergroup relations. Then, I present the dual-learning approach to attitude formation, in which both a propositional and an associative route are thought to underlie EC. I will explain how these routes are defined and that they can be tested by examining *when* EC occurs (i.e., operating conditions) and *what* type of representation is learned as an outcome of an EC procedure (i.e., operating principles).

The first two empirical chapters address (mainly) the operating conditions of EC. More specifically, we tested whether EC can be unconsciously (Chapters Two and Three) and efficiently (Chapter Two) obtained. We did so by relying on an attention load procedure in Chapter Two and on the evocativeness of the US in Chapter Three. The next two empirical chapters address (mainly) the operating principles of EC. Here, we tested whether the CS-US link encoded in memory is sensitive to a directional qualifier (Chapter Four) and if it reflects the statistical contingency between the CS and the US, and so extinguishes when the contingency decreases (Chapter Five).

In the General Discussion, I provide a summary of the various findings obtained and I discuss the theoretical implications of the empirical chapters. Here, I argue that neither 'hard' single-learning nor dual-learning approaches are the most viable and parsimonious approaches to EC. I propose an alternative, 'weak' single-learning approach to EC.

1. Evaluative Conditioning

Attitudes are "summary evaluations that we hold regarding objects, people and issues along a dimension ranging from negative to positive" (Petty, Wegener, & Fabrigar, 1997; p. 610). In principle, anything can be endowed with a certain valence (Osgood, Suci, & Tannenbaum, 1957). The construct of "attitudes" has long been at the core of the social psychological literature because of the assumption that a person's attitude is a major mediating variable between her exposure to information and her behavior. For instance, advertisement strategies are generally based on the idea that providing people with information about a given brand will result in the formation of favorable attitudes towards the brand, and so ultimately lead to buy products from this brand. In general, people approach what they like and avoid what they dislike, and this is true in many aspects of their life, including social interactions (Greenwald, McGhee, & Schwartz, 1998), voting behavior (Bassili, 1993, 1995), or food consumption (Hütter & Sweldens, 2018). Because attitudes are mainly acquired rather than inherited (Rozin, 1982; Rozin & Millman, 1987), it is crucial to understand how attitudes are formed and can be influenced.

Evaluative conditioning, as a CS-US pairing procedure, is generally considered one of the simplest approaches to attitude formation and change. Typically, CSs become positive after being paired with positive USs but negative after being paired with negative USs. EC, as a learning procedure, is straightforward and highly controlled, and it has been used in a variety of domains such as marketing (Sweldens, Osselaer, & Janiszewski, 2010), politics (Razran, 1954) and clinical interventions (Baccus, Baldwin, & Packer, 2004), just to name a few.

Research on EC can be traced back to research on classical conditioning. Pavlov (1927) noticed an unexpected phenomenon by observing the gastric functions of dogs: the repeated pairing of a stimulus (neutral stimulus: the sound of a bell) with another stimulus (the unconditioned stimulus: food) that is unconditionally linked to a response (the unconditioned response: salivation), leads the neutral stimulus to cause – by itself – the response (conditional response). This effect has inspired many studies on attitude formation wherein the repeated pairing between a neutral stimulus and a second, positive or negative, stimulus creates a positive or negative response to the – initially – neutral stimulus.

Staats and Staats (1958) were the first to examine whether "attitudes [could be] established by classical conditioning". They paired each of two national names (e.g., "Swedish" and "Dutch") with either 18 positive or 18 negative words by having their students read the national names (i.e., the CSs) as they appeared on a screen while the experimenter read aloud the affective words (i.e., the USs). The nation paired with positive words was evaluated later more favorably than the one paired with negative words.

In another pioneer study, Levey & Martin (1975) attempted to determine whether the evaluation of an initially neutral stimulus could be conditioned with stimuli that do not elicit a particular physiological response but are rather liked or disliked. In their study, the authors asked their participants to evaluate a series of postcards. The neutral cards were then paired with cards that had been evaluated either the most positively or the most negatively. These neutral-affective pairs were visually presented repeatedly. Following this conditioning phase, the initially neutral cards had acquired the valence of the postcards with which they had been paired.

Although early scholars referred to this phenomenon as “classical conditioning of attitudes” (e.g., Levey & Martin, 1975; Razran, 1954; Staats & Staats, 1958), this form of conditioning quickly earned its distinctive term of “Evaluative Conditioning” (Martin & Levey, 1978). EC may indeed be characterized by functional characteristics and reflect learning mechanisms that differ to those involved in classical conditioning (a question that is still debated to date; see also Chapter Five).

Since the pioneer study by Staats and Staats, research on EC has accumulated and is still very much alive today. A search on *PsychInfo* (in April 2018) indicates that there are 392 peer-reviewed scientific publications between 1978 and 2018 containing “Evaluative conditioning” in the title or the abstract. Part of this success relates to fundamental questions about the processes underlying EC, and in particular the possibility that it is driven by associative learning processes. This question will be at the core of this dissertation.

Other reasons relate to the flexibility and robustness of EC. Stimuli may be presented in visual (e.g., Levey & Martin, 1975), auditory (e.g., Gorn, 1982), olfactive (e.g., Davies, El-Deredy, Zandstra, & Blanchette, 2012) or haptic (e.g., Hammerl & Grabitz, 2000) modalities, and are not required to be both presented in the same sensory modality (e.g., Razran, 1938; Staats & Staats, 1958). Stimuli may be presented simultaneously or sequentially (for within-experiment manipulations, see Gawronski, Gast, & De Houwer, 2015; Hütter & Sweldens, 2013; Sweldens et al., 2010). EC has been observed both for backward pairings and forward pairings (for within-experiment manipulation, see Kim, Sweldens, & Hütter, 2015). EC is a robust and reliable effect of medium effect size (Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010).

Research on EC has predominantly focused on the procedural configurations under which the pairing of stimuli results in the formation of an attitude, although it has also recently started investigating the role of inter-individual differences in EC (Kendrick & Olson, 2012; Vogel, Hütter, & Gebauer, 2017). It is now well established that it is possible to influence people’s attitudes toward objects, (e.g., Stuart, Shimp, & Engle, 1987), individuals (e.g., Berkowitz & Knurek, 1969) and political slogans (e.g., Razran, 1938) by pairing them with affective stimuli in an EC procedure (see De Houwer, 2007; De Houwer et al., 2001; Hofmann et al., 2010; Walther, Nagengast, & Trasselli, 2005) for reviews). Procedurally, EC only requires a concomitant CS-US presentation.

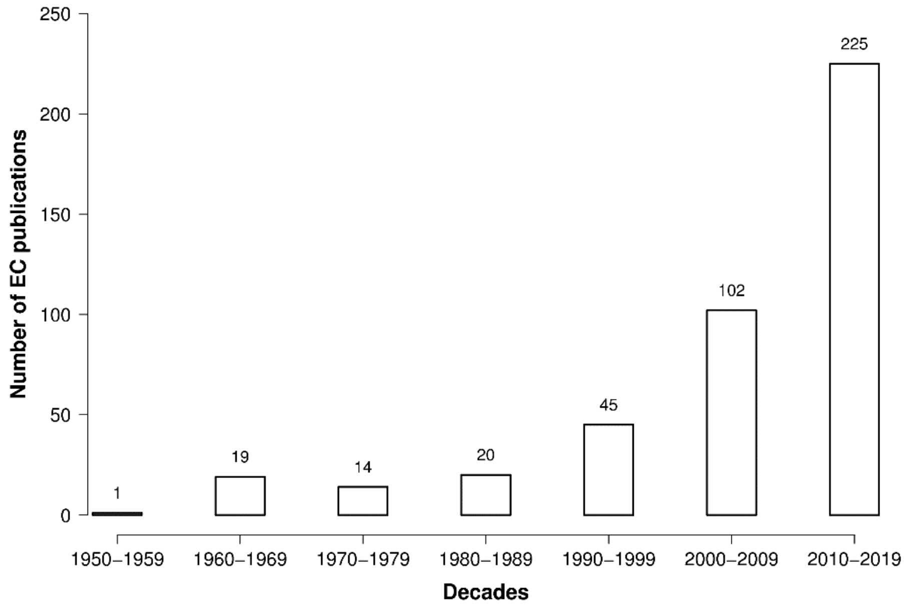


Figure 1. Number of peer-reviewed publications containing "Evaluative Conditioning" in the title and/or the abstract by decade. Data come from a *Psychinfo* search and were collected the 18th of April, 2018.

As just noted, the ever-growing body of research on EC is likely due to its wide applied potentials and high fundamental significance (Walther et al., 2005). Rozin, Wrzesniewski, and Byrnes, (1998) considered EC as "the main component of the acquisition of preferences". Below, I describe the applied and theoretical significance of EC.

2. Significance of Evaluative Conditioning

> "*Evaluative conditioning (EC) may well be pervasive in humans, and is currently the only articulated, experimental model of how likes and dislikes are created*." (Rozin et al., 1998, p. 398).

2.1 Applied significance of Evaluative Conditioning

By influencing attitudes toward virtually any kind of object, EC has applied potential in a broad variety of domains. EC may be successfully applied to marketing, politics, clinical psychology and intergroup relations. Below, I provide instances of classic EC studies conducted in these domains.

2.1.1 *Commercial marketing*

Typically, commercials about a given product (or brand) contain product-specific information but also background features such as pleasant colors, pleasant music, humor or the presence of celebrities or pleasant animals. These features do not necessarily contain any indication of the quality of the product (or brand). Rather, these affective stimuli – the USs – are used to elicit a positive reaction in the consumer, one that it is assumed to transfer to the target product (or brand). EC effects have been relied on for a long time in advertising.

> *"the ability of commercials to form associations by classical conditioning is well established and widely used"* (Hawkins, Best, & Coney, 1983, p. 314).

> *“To a large extent the consumer responds automatically, according to biologically determined patterns of behavior. Almost like an animal, s/he can be manipulated by classical conditioning. In short, there are biological limits imposed on consumer sovereignty-on the person's deliberate and conscious control of his/her behavior.”* (Kroeber-Riel, 1979, p. 240).

In the pioneer research of Gorn (1982), participants were exposed to pens of different colors (beige and blue) that were either presented while a pleasant music (from the movie “Grease”) or an unpleasant one (an Indian music). To thank the students for their participation, the experimenter told them they could take away one of the two pens of their choice as a gift. About 80% of the participants picked the pen associated with the liked music.

Many studies in consumer research established the influence of EC in promoting favorable brand attitudes (Allen & Janiszewski, 1989; Allen & Madden, 1985; Bierley, McSweeney, & Vannieuwkerk, 1985; Gibson, 2008; Gorn, 1982; Hütter & Sweldens, 2018; Janiszewski & Warlop, 1993; Kellaris & Cox, 1989; Kim, Allen, & Kardes, 1996; McSweeney & Bierley, 1984; Nord & Peter, 1980; Shimp, Stuart, & Engle, 1991; Stuart et al., 1987; Sweldens et al., 2010), even for changing the evaluation of “mature brands” such as *Coke* or *Pepsi* (Gibson, 2008) and to affect actual consumption behavior (e.g., Gibson, 2008; Gorn, 1982; Hütter & Sweldens, 2018).

2.1.2 *Socio-politic marketing*

Using a “luncheon technique”, Razran (1938) examined whether attitude towards political slogans may be influenced by pairing them with a positive or negative olfactory stimulation. Razran presented to students several political slogans (e.g., “America for Americans”, “Workers of the world unite”) either with a free lunch or with unpleasant odors. After pairing the slogans (i.e., the CSs) with the affective events (i.e., the USs) several days in a row, he asked his students to report their adherence to the political slogans (this time presented without lunch or odors). In what may be described as the first observation of an EC effect, Razran observed a stronger adherence to the slogans that were paired with positive USs as compared to the slogans paired with negative USs.

In the context of the US Presidential Elections of 1984, Mullen and colleagues (1986) examined whether “a smile can elect a president” (Mullen et al., 1986, p. 291). They observed that the presenter of *ABC News* smiled significantly more often when interacting with Ronald Reagan than with his adversary Walter Mondale. Presenters of other TV channels such as *CBS* and *NBC* did not show this bias. The authors observed that electors who watched *ABC News* more frequently than *CBS* and *NBC* were more likely to vote for Ronald Reagan than for Walter Mondale. The authors excluded the possibility that Reagan’s supporters were *a priori* more likely to watch *ABC News*.

In another US Presidential Election context, opposing George W. Bush to Al Gore, what was seen as an intentional use of “subliminal” EC for influencing citizens’ attitudes toward a political candidate was massively covered by the press. During a TV ad sponsored by the National Republican Committee, after displaying the image of Al Gore and just before the slogan "Bureaucrats decide" appeared for several seconds on the screen, a fragment of this word, "RATS", was flashed at the center of the screen for approximately 30ms. This ad was removed after it was displayed over 4000 times. In a fictional campaign spot, Weinberger and Westen (2008) observed that the brief presentation of the word “RATS” was indeed likely to increase negative judgments towards the target candidate.

It is common to see political figures surrounded by (positive) celebrities in public settings. These associations may also very well influence voters’ attitudes toward the political figure.

2.1.3 *Alcohol-dependence*

Research on addiction (Stacy & Wiers, 2010; Wiers & Stacy, 2006) suggests that an effective way to reduce drinking behavior is to change attitudes toward alcohol and that “a prime candidate for changing the affective value associated with alcohol is EC” (Houben, Schoenmakers, & Wiers, 2010, p. 1161). Houben and colleagues (Houben, Havermans, & Wiers, 2010; Houben, Schoenmakers, et al., 2010) reasoned that a brief intervention based on EC should change the attitude toward alcohol and, more importantly, reduce drinking behavior.

In one experiment, Houben and colleagues (2010a) manipulated the pairing of beer-related pictures (CSs) with or without negative words and pictures (USs). Next to this EC procedure, they measured in their participants beer-related attitudes, craving for beer, and actual drinking behavior both during a blind taste test and during the week that followed the intervention. They observed that participants for whom alcohol was paired with negative stimuli later exhibited both more negative beer-related attitudes and reduced drinking behavior.

These findings were also observed in two other experiments in which the EC procedure was less salient (taking the form of a video game or an advertisement, (Houben, Schoenmakers, et al., 2010) and were recently replicated in an independent pre-registered experiment (Tello, Bocage-Barthélémy, Dandaba, Jaafari, & Chatard, 2018). In the latter experiment, participants who received the intervention reduced on average their weekly consumption of alcoholic beverages by 31% (as compared to 4% in the control condition). The authors further observed that the effect of the EC procedure on drinking behavior benefited to a similar extent to both control – healthy – subjects and hazardous drinkers (hazardous drinkers have a pattern of alcohol consumption that places them at risk for adverse health events).

Taken together, these results suggest that a brief EC intervention may have clinical implications for reducing drinking behavior, which is especially important for risky populations such as hazardous drinkers.

2.1.4 *Self-esteem*

Self-esteem is defined as “the individual’s positive or negative attitude toward the self as a totality” (Rosenberg, Schooler, Schoenbach, & Rosenberg, 1995, p. 141). A large-scale study suggests that low self-esteem has a host of negative consequences, including substance abuse and hostile behavior (Parker & Benson, 2004, see also Harter, 1993; Leary, Schreindorfer, & Haupt, 1995).

Self-esteem has also been described as “the association of the concept of self with a valent attribute” (Greenwald et al., 2002, p. 5). As such, one might consider using an EC procedure to improve self-esteem. In a series of 6 experiments, Dijksterhuis (2004) proposed to conceptualize self-esteem as associative links between the concept of self and positive or negative information. Consistently, the author showed that participants reported having a higher self-esteem when stimuli representing the self (the CS) were paired with positive words (the USs) than when not. As discussed elsewhere (Pleyers, Corneille, Luminet, & Yzerbyt, 2007; Stahl, Haaf, & Corneille, 2016), however, the design of the experiments did not allow to properly conclude that the pairing between the self and the USs led to a change in self-esteem. This is because conditions either differed in the presence of positive words (potentially leading to a difference in mood) or self-related words (potentially leading to a difference in the self-salience).

Baccus and colleagues (2004) tested the same hypothesis but relied on a counterbalanced design in which both self-related and positive words appeared in a similar frequency for both experimental conditions. They used as CSs both self- and other-related words, which were either paired with positive or negative faces. They observed that participants' self-esteem became more positive after the self was paired with positive than negative USs. Of note, the influence of the EC procedure on self-esteem was observed with an indirect self-esteem measure only (i.e., an Implicit Association Test, or IAT; Greenwald et al., 1998)[1].

These studies suggest that self-esteem may be sensitive to a brief EC intervention and that even a CS that is not initially neutral and unfamiliar (i.e., the self) can be sensitive to an EC procedure.

2.1.5 *Prejudice*

Prejudice relates to negative affective reactions about social groups (e.g., racism, sexism). Prejudice is a main predictor of discriminatory behavior (e.g., Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; McConnell & Leibold, 2001). It has been suggested that being repeatedly exposed to the pairing of members of social groups and negative events leads to the development of prejudice towards these groups (e.g., De Houwer et al., 2001; Olson & Fazio, 2001).

Olson and Fazio (2006) reasoned that if prejudice is learned through associations, it could as well be reduced using an EC procedure. In three experiments, they paired pictures of Afro-American and Caucasian individuals and positive and negative pictures (Afro-American-positive and Caucasian-negative pairings). They observed that the conditioning procedure reduced prejudice as measured with an indirect evaluative measure, and this reduction was still observed at least two days after the conditioning procedure.

2.2 Theoretical significance of Evaluative Conditioning

Attitudes were originally seen as emerging from the careful weighting of pro- and anti-attitudinal information towards a given attitudinal object (e.g., Hovland, Janis, & Kelley, 1953; Sherif & Hovland, 1961). This conception was quickly followed by a dual-process approach, according to which attitudes originate from two distinct learning systems or "routes" (Chaiken, Liberman, & Eagly, 1989; Gawronski & Bodenhausen, 2006; Petty & Cacioppo, 1986). A first system – cognitively elaborated – is typically thought of as involving attention, understanding and acceptance of a message and leading to an informed attitude. A second system – cognitively automatic – is typically thought of as involving simple affective processes that lead to an incremental change in an "associative structure", the latter of which may be described as "a network of variable-strength associations" (Greenwald et al., 2002, p. 5).

[1] Attitudes inferred using indirect measures such as the IAT may be called "implicit attitudes" although it does not necessarily mean they reflect constructs distinct from "explicit attitudes" as obtained via self-reported measures.

As highlighted by Sweldens, Corneille, and Yzerbyt (2014), "The fact that attitudes could be changed through simple conditioning procedures was one important reason to posit the existence of a qualitatively different peripheral processing route, given that attitudinal conditioning was considered to invoke minimal processing resources and operate without awareness." (Sweldens et al. 2014, p. 189).

> "At the low end of the elaboration continuum (…) attitude change can result from a number of less resource-demanding processes, such as classical conditioning." (Petty & Wegener, 1999, p. 42).
>
> "The prototypical case for implicit attitude changes resulting from changes in associative structure is evaluative conditioning" (Gawronski & Bodenhausen, 2006, p. 697).

EC fits well with the notion that associations between two concepts are formed and strengthened through their simultaneous activation (Greenwald et al., 2002; Hebb, 1949). The contiguous presentation of a CS with a US would lead to an associative link between the CS and the US in the associative structure. At first sight, EC is therefore an interesting candidate for an associative learning route. However, as discussed elsewhere (Corneille & Stahl, 2018; Sweldens, et al., 2014), the claim that EC occurs automatically has been heavily questioned, shedding doubts on the theoretical value of the dual-learning approach to attitude formation, as compared to more parsimonious "single-process" approaches (De Houwer, 2009; Kruglanski & Gigerenzer, 2011; Mitchell, De Houwer, & Lovibond, 2009; Newell & Shanks, 2014).

In the next section, I discuss how "associative" learning is defined in the dual-process framework and how it can be tested in the context of EC studies. We will see that inferring the activity of an associative learning process in EC implies to examine the boundary conditions of EC (i.e., its operating conditions) and the nature of the evaluative representation learned in an EC procedure (i.e., its operating principles; Corneille & Stahl, 2018; Gawronski & Bodenhausen, 2009, 2014; Sweldens, 2018).

Throughout this dissertation, the terms "automatic", "associative" and "implicit" will be used. They are often used interchangeably in the literature but are best described differently. Although these constructs have been (and still are) largely debated, we may sketch their commonly accepted meaning as follows.

The term "automatic" usually refers to the dependency of a given process on four criteria: awareness, efficiency, controllability, and intentionality (Bargh, 1994). Determining the automaticity of a process is difficult for at least two reasons. First, the four criteria can have many different definitions (see Moors & De Houwer, 2006, for an in-depth discussion). Second, the criteria do not necessarily perfectly align (Bargh, 1994; Moors & De Houwer, 2006). In theory, by considering four sets of binary features, one could expect $2^4 = 16$ categories of processes (Melnikoff & Bargh, 2018). I will describe more precisely each of the four automatic features as well as their role in EC in a following section.

The term "associative learning" has two meanings, which are often confused. The first refers to the mere learning that two (or more) events in the world are related to another. The second meaning specifies the underlying mental mechanism and conditions of their occurrence. More specifically, the term "associative" refers here to (1) the automaticity of learning through the perception of co-occurrences in the environment, resulting in (2) the formation of excitatory or inhibitory links between co-occurring events represented in memory (e.g., Gawronski & Bodenhausen, 2014).

The term "implicit" covers "what is not explicit". It also has two meanings that are often confused with each other. One meaning refers to the construct and the other to the measurement. First, an implicit construct can be defined as "the introspectively unidentified (or inaccurately identified) trace of past experience that mediates the relevant category of responses" (Greenwald & Banaji, 1995, p. 5). It therefore stresses the idea that the construct should not necessarily be accessible to awareness, endorsed, or intentionally used to influence a given response. Second, the term "implicit" refers to indirect measures used to objectify this construct, such as the Implicit Association Test (IAT, Greenwald, McGhee, & Schwartz, 1998).

3. Dual-learning approach to Evaluative Conditioning

Dual-learning models suggest that EC has two qualitatively distinct sources: it can be established through propositional and associative learning processes.

One of the most illustrative experiments claiming support for dual-process models was published by Rydell, McConnell, Mackie, and Strain (2006). These authors reported that explicit attitudes (i.e., obtained with a direct evaluative measure: a Likert scale) and implicit attitudes (obtained with an indirect evaluative measure: an IAT) toward an individual (Bob) are influenced by supraliminal and subliminal valent information, respectively. In their procedure, a picture of Bob was displayed after a rapid presentation (i.e., 25 ms) of valent words (e.g., love or death) and was then directly followed by the supraliminal presentation of the description of a positive or negative behavior characterizing Bob. Importantly, the subliminal and the supraliminal USs were of opposite valence. For instance, Bob could be presented with "death" briefly flashed on the screen and with "helps the old lady" as presented supraliminally. Their results showed a perfect double-dissociation between direct and indirect evaluations. Whereas participants rated Bob positively on the self-reported measures (i.e., positive explicit attitude), they were more prompt to categorize negative words along with Bob (i.e., negative implicit attitude).

Although this experiment has not been replicated (Heycke, Gehrmann, Haaf, & Stahl, 2018; Rosocha & Balas, 2017), its underlying *rationale* illustrates two major assumptions that can be found in the dual-process approach: (1) there are two independent learning processes, (2) that are sensitive to distinct operating conditions (i.e., conscious-dependent and independent). Many dual-learning models, such as the Systems of Evaluation Model (or SEM, McConnell & Rydell, 2014), go the extra step and assume that propositional and associative learning preferentially impact direct and indirect evaluative measures, respectively. I use here the term "dual-process approach" because there is not a unique dual-process theory but rather different dual-learning theories that share common assumptions (e.g., Jones, Olson, & Fazio, 2010; Sternberg & McClelland, 2012; Sweldens et al., 2010). They all share the assumptions that EC can be driven by propositional and by associative processes. In the propositional route, EC is thought to reflect propositions about the CS-US relations that were non-automatically formed through a propositional and rule-based process. In the associative route, EC is thought to reflect CS-US associations that were created by associative processes that automatically registered the mere CS-US contiguity. Since they cannot be directly observed, the processes underlying EC have to be inferred from their operating conditions and operating principles (Corneille & Stahl, 2018; Gawronski & Bodenhausen, 2014). In this regard, associative and propositional processes are typically defined in opposition.

Below, I discuss the operating conditions and then operating principles representative of associative learning, some of which will be addressed in the empirical chapters.

3.1 Operating conditions

In contrast to the formation of propositions, the formation of associations is thought to be automatic. The "automaticity" of the process refers to its (in)dependence from (1) contingency awareness, (2) attentional resource, (3) processing goals and (4) intentional control (Bargh, 1994; Moors & De Houwer, 2006). These "operating conditions" (Gawronski & Bodenhausen, 2009; 2014) describe *when* EC is expected to occur.

If EC reflects a propositional process it is expected to occur only when (1) the subject is aware of the CS-US contingencies, (2) the subject has sufficient available attentional resources, (3) the subject has related processing goals and (4) the subject does not intent to prevent the process to occur (De Houwer, 2009). If EC reflects (or can reflect) an associative process, at least some manifestations of EC can occur automatically. For instance, EC may be observed when the awareness of the CS-US contingencies is prevented (through the subliminal presentation of the CS, for instance).

An associative process requires very minimal conditions to occur. Early associative accounts of EC mentioned the necessity that the US elicits an evaluative response and the perception of a contiguity between the CS and the US (Frank Baeyens, Crombez, Van den Bergh, & Eelen, 1988; Levey & Martin, 1975; Martin & Levey, 1978). More contemporary views of associative processes in EC, however, assume additional enabling conditions.

More specifically, an associative process taking the form of an "implicit misattribution" of the affect, in which the evaluative response elicited by the US is wrongly attributed to the CS, requires the simultaneous processing of the CS and the US and benefits from enhancing confusion about the source of the evaluative response (Jones, Fazio, & Olson, 2009; Jones et al., 2010). Relatedly, a "direct affective transfer" of the evaluative response elicited by the US to the CS is thought to benefit from disposing of "one-to-many" CS-US pairings in which a given CS is paired with multiple USs sharing the same valence.

Obtaining evidence for an automatic (e.g., unaware or efficient) EC would be consistent with the dual-process view as it would suggest that associative learning can underlie EC. Not obtaining an EC effect in settings preventing a propositional process (e.g., under a subliminal presentation of the CS) does not however reject the dual-process approach in a straightforward manner. Indeed, one must pay attention to the settings enabling or facilitating an associative process, such as (according to some associative accounts) simultaneous and one-to-many CS-US pairings.

3.2 Operating principles

Whereas operating conditions are concerned with conditions under which EC can be observed, operating principles are concerned with how EC is operating and what representation it creates in memory. In contrast to associations, propositions are truth-sensitive (they can be rejected or not) and represent relations between the CS and the US (De Houwer, 2009; Mitchell et al., 2009). As such, propositions represent *qualified* linkages between the CS and the US whereas associations are *unqualified*. These "operating principles" address *what* is learned in EC (Gawronski & Bodenhausen, 2009; 2014); that is, representations that passively register mere co-occurrences between the CS and the US *versus* that register their meaning (i.e., their truth or relation).

The "unqualified" CS-US linkage refers to the idea that the link stored in memory between the CS and the US is merely excitatory: the activation of one stimulus causes the activation of another associated stimulus (e.g., Dickinson, 1980; Shanks, 2007). When the CS is activated, it would automatically activate the response linked with the US either directly (stimulus-response association) or indirectly by activating the US mental representation (stimulus-stimulus association). In contrast, propositions contain information that qualifies how events are related. The propositions "A causes B" and "B causes A" are two different propositions although only one association links A and B. This qualification implies for instance that one could expect reversed EC effect, i.e., a change in valence of the CS in opposition to its paired US. For instance, if one character stops an unpleasant sound, it should be liked although it is contiguous with a negative US.

Furthermore, associations do not contain truth information (e.g., Strack & Deutsch, 2004): "associations are neither true nor untrue, they just are." (De Houwer, 2009, p.3). In contrast, propositions can be endorsed or not, they can differ in the extent to which they are believed to be accurate. This “truth validation” depends on how the new proposition fits with previously encoded ones (e.g., Gawronski & Strack, 2004).

Hence, just as for the operating conditions, testing the sensitivity of EC to operating principles sheds light on the associative versus propositional nature of its underlying processes. If EC does not depend on the way CS and US are related, but merely depends on their mere contiguity, at least in some circumstances, it would be reasonable to consider that associative processes contribute to EC.

To date, there is little doubt that EC can be driven by propositional process, but what remains to be demonstrated is which operating conditions and which operating principles are predictive of EC (Corneille & Stahl, 2018; Sweldens et al., 2014). These questions motivated the research presented in this dissertation. Before introducing the empirical chapters, I will briefly summarize evidence relevant to associative learning in EC, by discussing separately research on the automaticity of EC and research on its qualification.

4. When does the EC effect occur (operating conditions)?

A first way to infer process(es) underlying EC is to test its boundary conditions, or “enabling conditions” (De Houwer, 2007), or “operating conditions” (Corneille & Stahl, 2018; Gawronski & Bodenhausen, 2009). These operating conditions are (1) awareness of the CS-US contingencies, (2) the availability of attentional resources, (3) the presence of certain goals, and (4) the intentional control. Investigating these operating conditions relates to whether EC can be elicited automatically (Bargh, 1994; Moors & De Houwer, 2006). Evidence for automatic EC would support dual-process models of attitude.

4.1 Contingency awareness

The role of contingency awareness in EC has been extensively addressed. It is viewed as of “paramount theoretical importance (…) for attitude formation theories specifically and human learning theories in general” (Sweldens et al., 2014, p. 188) and it motivated about 25% of the research on EC (Sweldens et al., 2014).

> “The real debate is whether or not humans possess a learning mechanism that allows direct affect transfer between paired stimuli as part of an automatic, unconscious process.” (Allen & Madden, 1985)

This question has been extensively discussed elsewhere (see Corneille & Stahl, 2018; Sweldens et al., 2014, for recent reviews) and will therefore not be reviewed in details in this dissertation. I will rather sketch the current evidence and open questions.

It is important to highlight from the outset that the study of contingency awareness in EC is characterized by a fierce debate: “It seems that for every article claiming that EC procedures can change attitudes without participants' awareness (…) there is an article claiming that EC does not happen without awareness" (Sweldens et al., 2014, p. 188). After 40 years of research on EC, one should remain cautious about claiming the (in)dependence of EC to the CS-US contingencies awareness.

A first reason for the current uncertainty relates to the very definition of "awareness". Sweldens and colleagues (2014) called for attention to the lack of definitional clarity of contingency awareness. Originally, researchers were interested in the awareness of the impact of the pairings on the resulting evaluations (e.g., Allen & Janiszewski, 1989; Staats & Staats, 1957). Then, it became progressively clear that the awareness of the CS-US contingencies represented a more stringent criterion. Researchers have relied on experimental and correlational methods for studying the role of CS-US contingency in EC. The experimental method consists in preventing the encoding of the CS-US pairings, and its success is probed using CS-US identification tasks at a later stage. The correlational method relies on the latter memory measures only; that is, in the absence of experimental induction interfering with CS-US encoding.

The "awareness" addressed in the experimental studies is certainly closer to the definition of "conscious perception" (Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006), whereas the "awareness" addressed in correlational studies is closer to the definition of episodic memory (Tulving, 1972). Whereas there is little evidence for unaware EC in studies manipulating contingency awareness, there are more conflicting results in studies using a correlational design (Sweldens et al., 2014). Below, we provide a summary discussion of the experimental and correlational approaches to the role of awareness in EC.

When contingency awareness is prevented by presenting the CS very briefly (Stahl et al., 2016), or parafoveally (Dedonder, Corneille, Bertinchamps, & Yzerbyt, 2014) or when suppressing it through a Continuous Flash Suppression procedure (Högden, Hütter, & Unkelbach, 2018) no EC effect is observed. In two unpublished experiments (Mierop, Corneille, Stahl, & Vermeulen, *unpublished*), we adapted the attentional blink paradigm (Duncan, Ward, & Shapiro, 1994) and developed a new EC paradigm designed to manipulate awareness independently of attentional resources and perception. The attentional blink paradigm uses rapid sequential presentation of stimuli during which targets (e.g., words) are embedded among distractors. Participants are typically asked to report the two targets right after the presentation of the stream. When the two targets are presented with a short SOA (e.g., 300 ms), the semantic representation of the second target (T2) is activated but without accessing awareness. This phenomenon is called "attentional blink" (Raymond et al., 1995). In our experiments, T1s (CSs) were paired with either positive or negative T2s (USs). In order to examine the moderating role of awareness in EC, we manipulated the SOA on a within-subject basis, such that some CSs were paired with blinked USs (shorter SOA trials) whereas other CSs were paired with non-blinked USs (longer SOA trials). The rationale was that if EC occurs independently from contingency awareness, CS evaluations should only depend on US valence and should not be affected by SOA duration. Alternatively, if EC requires CS-US contingencies awareness, EC should only occur for the longer SOAs (i.e., for non-blinked USs).

We obtained a strong and reliable attentional blink effect although EC was, on average, not observed and was not moderated by SOA. When looking at the EC dependency on the proportion of correct identification, EC was only observed for the higher levels of awareness (i.e., when the USs were never blinked for a given CS).

Turning to correlational studies, conclusions are less straightforward. Whereas experiments relying on an item-based analysis of the contingence memory reported EC only for pairs correctly reported (Pleyers et al., 2007; Stahl & Unkelbach, 2009; Stahl, Unkelbach, & Corneille, 2009), experiments relying on a purer Process Dissociation (or PD) procedure (which is considered the "best [correlational] approach to study whether attitudes can be conditioned without contingency awareness", Sweldens et al., 2014, p. 196) reported EC without explicit memory of the pairings (Hütter & Sweldens, 2013; Hütter, Sweldens, Stahl, Unkelbach, & Klauer, 2012; Mierop, Hütter, Stahl, & Corneille, 2018). I now develop the underlying rationale of the PD task used in EC research, as it is critical to the first two empirical chapters of this dissertation.

The PD approach was designed to separate the contributions of two independent processes in reporting the US valence attached to a particular CS: memory of the US valence versus affect-based inferences about US valence. Hütter and colleagues (2012) argued that when providing a memory judgment about the US valence (e.g., "was this CS paired with positive or negative USs?") one could use either one's memory of it ("I remember this CS was paired with positive USs") or infer the US valence on the basis of the CS evaluation ("I like the CS so I infer it was paired with positive USs"). The use of the latter affect-as-heuristic process suggests that an attitude has been learned and applied although the pairings is not retrieved. This may lead to the overcategorization of CS-US pairs as "aware" and therefore hide "unaware" pairs and consequently "unaware" EC.

The rationale of the PD consists in contrasting responses under conditions where the two above processes operate in the same direction (inclusion condition) versus in antagonistic directions (exclusion condition). In the inclusion condition, participants are instructed to respond "pleasant" ("unpleasant") if they remember that a CS was paired with positive (negative) USs. Importantly, they have to draw on their attitude when they do not remember the pairings. Specifically, they are instructed to select "pleasant" ("unpleasant") for liked (disliked) CSs. In this inclusion condition, contingency memory and attitude without memory elicit the same answer. In contrast, in the exclusion condition, participants are instructed to reverse their memory-based response (or, in a variant of the task, their attitude-based response). This time, the two processes are now antagonistic to each other.

By comparing the frequencies of the “pleasant” and “unpleasant” responses between the inclusion and the exclusion condition in a Multinomial Processing Tree model, one can estimate parameters that are supposed to reflect the contribution of these two processes. The first, M, parameter is thought to reflect the contribution of actual contingency memory in the responses. The second, A, parameter is thought to reflect the contribution of the affect-as-heuristic in the responses. A third, R, parameter is thought to reflect a response bias towards the response “pleasant”.

The model fit assesses whether the observed response frequencies deviate from the response frequencies predicted by the model (assuming the contribution of both explicit contingency memory and affect-as-heuristic). The crucial test in the context of the automaticity of EC is the test of the A parameter against 0. If fixing the A parameter to 0 significantly reduces the model fit, then one may conclude that the affect-as-heuristic influenced substantially the responses in the memory task. Consequently, one is tempted to conclude that attitudes may be learned (or at least maintained) without explicit memory for the CS-US pairs.

Importantly, the only test of the model fit does not constitute *per se* sufficient evidence for the validity of the parameters (Hütter & Klauer, 2016). In addition, one should test the construct validity of the parameters by demonstrating they capture the processes they are intended to. In this perspective, in their fifth experiment, Hütter and colleagues (2012) observed that the a-parameter is resistant to a time delay of 24 hours whereas the M-parameter is not, “confirming the dissociability of the attitude and memory components.” (Hütter et al., 2012, p. 539). In another experiment, Hütter and Sweldens (2013) predicted and observed the a-parameter only for simultaneous CS-US pairings. As explained below, this finding was consistent with dual-process models hypothesizing that “unaware EC effects are caused by implicit misattribution, which is only possible with simultaneous CS–US presentations. (Hütter & Sweldens, 2013, p. 4). Of note, not only the a-parameter but also the m-parameter was reduced (though not to non-significance) in the sequential pairing condition.

As we will see in the first two empirical chapters, the a-parameter has however not always been replicated in procedural settings “conducive” to an implicit misattribution process (Mierop, Hütter, & Corneille, 2017) and it does not seem to depend on another crucial implicit misattribution factor; namely, the affective evocativeness of the US (Mierop et al., 2018). Furthermore, this parameter is observed when participants are merely instructed about CS-US pairings even though no actual CS-US pairings are presented to them, and it is sensitive to retrieval strategies (Hütter & De Houwer, 2017). Overall, it is therefore safer to conclude that the a-parameter reflects an EC effect without strong contingency memory, rather than an EC effect resulting from an unconscious associative encoding process.

As we just discussed, uncertainties regarding the role of awareness in EC are largely dependent on the diversity of operationalization of this construct. Overall, more stringent approaches (i.e., experimental approaches relying on item-based analyses) provide little support to the unaware EC view. A second reason we just briefly alluded to, however, is that different EC paradigms may be conducive to distinct processes that are differently related to awareness (Jones et al., 2010; Sweldens et al., 2014). In line with this dual-process approach, several experiments claiming unaware EC (more precisely: EC observed without explicit memory of the pairings) involved paradigms that were conducive to an "implicit misattribution process" or "direct affective transfer" (i.e., one-to-many simultaneous pairings) but not in paradigms preventing the latter associative processes (i.e., sequential CS-US pairings; Hütter & Sweldens, 2013; Jones et al., 2009; Sweldens et al., 2010). It has therefore been suggested that: "(…) a simultaneous presentation format can be critical in producing EC effects under conditions that compromise the operation of propositional processes." (Hütter & Sweldens, 2013, p. 4).

As we will see in the first two empirical chapters, however, unaware EC (still defined as EC without memory of the pairings) is not always observed in these settings (Mierop, Hütter, & Corneille, 2017; Chapter Two) and critical factors that are thought to enable or at least to facilitate associative learning, such as the US valence evocativeness (which refers to the affective intensity and its processing speed), do not reliably predict the occurrence of unaware EC (Mierop et al., 2018; Chapter Three).

Overall, the evidence for unaware EC is rather scarce and since it is supported by correlational studies addressing the role of contingency awareness with memory tasks, one may question whether it reflects an associative learning process or the maintenance of EC despite a contingency memory decay (I develop this point in the General Discussion). The conditions for observing an unaware EC have yet to be more thoroughly understood.

4.2 Attentional resources

Another crucial operating condition of EC is the availability of attentional resources. Automatic processes can occur even with minimal attentional resources (Bargh, 1994). If EC relies (or can rely) on an automatic encoding process, one should expect to observe EC even when attentional resources are depleted at encoding.

The first experiments relying on a manipulation of attentional resources are difficult to interpret because they either relied on very small sample (e.g., Fulcher & Hammerl, 2001), or on a manipulation not restricted to reducing the availability of attentional resources (e.g., Field & Moore, 2005; Fulcher & Hammerl, 2001) or unevenly depleting the resources throughout the exposition to the CS-US pairs (e.g., Walther, 2002).

Pleyers, Corneille, Yzerbyt, and Luminet (2009) compared a resource reduction condition in which participants were performing a two-back task (continuous retention task) to a control condition in which participants listened to neutral music. They observed that reducing attentional resources reduced both EC and contingency memory. These results where replicated when using novel and unfamiliar CSs (Dedonder, Corneille, Yzerbyt, & Kuppens, 2010) and when using a conditioning paradigm arguably conducive to Implicit Misattribution (Mierop et al., 2017; Chapter Two). One could argue that such distracting task might prevent the very visual processing of the stimuli (Blask, Walther, Halbeisen, & Weil, 2012). Contrary to this argument, in a follow-up study, Pleyers and colleagues (2009) observed that up to 76% of the individual stimuli presented in the resource depletion condition were correctly recognized. Consistent with these findings, Kattner (2012) showed that selectively diverting the attention from the contingency between stimuli prevents EC to occur even if the stimuli are individually attended. This author used two distinct distraction tasks to divert the attention from the contingencies (by asking subjects to learn irrelevant contingencies) or from the stimuli (by realizing math equations). These distraction tasks both reduced EC to non-significance and this detrimental effect was mediated by a reduction of CS-US contingency memory. This is consistent with the idea that top-down attention is required for conscious processing (Dehaene et al., 2006), which would then be required for EC.

In sum, it is reasonable to assume that the availability of attentional resources is necessary for consciously encoding the contingencies and, consequently, for the establishment of EC. So far, however, resources manipulations involved comparing two groups of participants, one of which was distracted. The problem with a between-subject design is that it may lead to different processing goals depending on the experimental condition. For instance, it is arguable that the “resource depletion” condition does not only deplete participants’ attentional resources but also distracts them from the evaluative processing of CS-US pairings. As discussed in the following section, EC is sensitive to processing goals, including goals related to the evaluative processing of CS-US pairs (Gast & Rothermund, 2011).

Social psychologists often separate the awareness and efficiency criteria although they seem largely related to one other. For instance, reducing the availability of attentional resources results in a weaker awareness of the pairings (e.g., Pleyers et al., 2009). According to Dehaene and colleagues (2006), the level of conscious perception depends on both bottom-up stimulus strength (weak vs. strong) and top-down attention allocated to the stimulus (absent vs. present). From the lower to the stronger state of consciousness, the perception can be subliminal (unattended), subliminal (attended), preconscious and conscious. As we saw, reducing the stimulus strength (e.g., Stahl et al., 2016) or the top-down attention allocated to the CS (e.g., Dedonder et al., 2013; Högden et al. 2018) or the CS-US pairings (e.g., Mierop et al., 2017) prevent the establishment of EC.

		Top-down attention	
		Absent	Present
Bottom-up stimulus strength	Weak	Subliminal (unattended)	Subliminal (attended)
	Strong	Preconscious	Conscious

4.3 Processing goals

An automatic process occurs regardless of the motivations and goals of the individual (Bargh, 1994). If EC can occur automatically, it should therefore not depend on goals.

Corneille, Yzerbyt, Pleyers, and Mussweiler (2009) manipulated the focus of their participants on similarities or differences between pairs of pictures. They asked half of their sample to identify a maximum of similarities between two drawings for two minutes; the other half had to identify a maximum of differences. After this manipulation of processing goals, participants realized the conditioning phase in which consumer products (CSs) were paired with positive or negative images (USs). Although EC was observed in both experimental conditions, the authors observed that it was stronger in the similarity condition. This beneficial effect of a similarity focus on EC was recently replicated

(Mierop, Hütter, Bret, & Corneille, *unpublished*). In this experiment, we used both direct and indirect evaluative measures to test whether the processing goals influences the formation of an attitude or rather its expression. As we will see in the following section, the main difference between explicit measures of attitudes (captured through direct evaluative measures) and implicit measures of attitudes (captured through indirect evaluative measures) is that the latter are less likely to reflect a “controlled” evaluation (Gawronski, Balas, & Creighton, 2014). We obtained a similar pattern of results on the explicit and implicit attitudes although the influence of the processing goals on the EC effect reached a conventional level of significance only on the direct measure.

Gast and Rothermund (2011) manipulated processing goals related to the evaluative processing rather than the relation between the stimuli. They hypothesized that EC should only occur when individuals direct their attention to the affective nature of the US. In their first study, they presented pairs of portraits to their participants. Each pair consisted of a neutral portrait (CS) and a portrait liked or not (US). They asked half of their participants to indicate the extent to which they liked portraits (affective focus) and the other half if they thought the people were from the south or the north of Germany (cognitive focus). The results indicated that EC occurred in the affective focus condition but not in the cognitive focus condition. The authors concluded that focusing on the valence of CS-US pairs is necessary for EC to occur.

Verwijmeren, Karremans, Stroebe, and Wigboldus (2012) manipulated a physiological goal and examined its influence on attitude formation for stimuli that are relevant to this goal. More specifically, they induced thirst in half of their participants before submitting them to a conditioning phase wherein neutral beverages (CSs) were paired with disgusted faces (relevant USs) or frightened faces (irrelevant USs). They observed that, in thirsty participants, EC was stronger for the CSs paired with relevant USs.

In summary, processing goals seem to influence EC. Focusing on the affective dimension of the stimuli (Gast & Rothermund, 2011), on the relationship linking the CS and US (Corneille et al., 2009) or on stimuli relevant to a physiologically-induced goal (Verwijmeren et al., 2012) favor the emergence of EC.

4.4 Controllability

An automatic process cannot be controlled: it is impossible to deliberately disrupt its activity (Bargh, 1994). Balas and Gawronski (2012) examined whether it is possible to deliberately reduce or increase EC. According to the authors, if EC does not occur automatically, it should be possible to control the expression of conditioned attitudes if the subjects are (1) motivated to follow the instructions to promote or prevent the impact of the CS-US pairings, (2) capable and (3) recognize that a given CS was paired with positive or negative US. In two experiments, they observed that instructing participants to maximize the influence of CS-US pairings on subsequent evaluations of the CSs (i.e., the EC effect) increased EC as compared to instructing them to minimize or reverse EC, but only when they remembered correctly the pairings.

These experiments suggest that the attitude expression is controllable but do not allow any conclusion on the controllability of attitude formation (Gast, Gawronski, & De Houwer, 2012). To distinguish between the impact of control on the formation versus expression of attitudes, (Gawronski et al., 2014) (1) instructed their participants to maximize or minimize the EC effect *before* being exposed to the CS-US pairings, and (2) used both direct and indirect evaluative measures, the latter of which were assumed to be uncontrollable. They observed EC on both type of evaluative measures but only the explicit attitude was influenced by the intentional control of the participants. They concluded that "that, although the *expression* of self-reported CS evaluations can be intentionally controlled, the *formation* of conditioned attitudes is rather difficult to control" (Gawronski et al., 2014, p. 428).

Recently, the uncontrollability of the EC effect has been replicated using different controllability manipulations (Gawronski, Gast, et al., 2015) and Multinomial Processing Tree Modeling (Hütter & Sweldens, 2018). The latter work is particularly convincing, the uncontrollability parameter is a valid predictor of behavior and was unaffected by experimental manipulations of motivation and attentional resource. It is therefore consistent with the view that EC, at least in some circumstances and to a certain extent, is driven by uncontrollable learning processes.

5. What is learned in an Evaluative Conditioning procedure (operating principles)?

In comparison to operating conditions, attention to the operating principles of EC has been rather neglected (Sweldens, 2018), although operating principles also allow testing the associative nature of the learning processes involved in EC.

In this section, I discuss evidence about what CS-US linkage is stored in memory. Two particularly relevant areas of research can be identified (Corneille & Stahl, 2018). The first one examined whether the CS-US linkage is qualified by relational and truth-validity information. The second one examined whether the CS mental representation is directly linked to the US response (stimulus-response linkage) or indirectly, through the mental representation of the US (stimulus-stimulus linkage).

5.1 Unqualified or qualified CS-US linkages

As discussed earlier, one way to determine whether associations or propositions are formed in memory in EC procedures is to examine the qualifications of the CS-US linkages. The main idea is that associations are unqualified CS-US linkages whereas propositions are qualified CS-US linkages (Corneille & Stahl, 2018; De Houwer, 2009; Gawronski &

Bodenhausen, 2009, 2014). Whereas associations capture only the strength of the relations between events, propositions additionally capture their structure: they include information about *how* the events are related (Lagnado, Waldmann, Hagmayer, & Sloman, 2007; Shanks, 2007).

In a pioneer study, Gawronski, Walther, and Blank (2005) presented pairs of faces to their participants. Half of these faces were neutral (CSs) whereas the other half were positively or negatively valent (USs). They further added the relational information “likes” or “dislikes” between the CS face and the US face to describe the feelings that one shares with the other. For the pairs of faces characterized with the “likes” feelings, they found EC both on direct and indirect evaluative measures. CS faces paired with positive US faces were preferred to CS faces paired with negative US faces. For the pairs characterized with “dislikes” feelings, they found reversed EC both on direct and on indirect evaluative measures. CS faces that disliked negative US faces were preferred to CS faces that disliked positive US faces. These results were replicated on direct evaluative measures with different relational information (e.g., “friend of” vs. “enemy of”, “loves” vs. “loathes”) manipulated between-participants (Fiedler & Unkelbach, 2011) or within-participants (Förderer & Unkelbach, 2012; Zanon, De Houwer, Gast, & Smith, 2014).

As compared to associations, propositions are also thought to be qualified by a truth validation (De Houwer, 2009; Mitchell et al., 2009). Hence, manipulating the validity of the CS-US pairings during an EC procedure should influence EC if it does reflect the formation of propositions. Peters and Gawronski (2011) manipulated the validity of the CS-US pairings by presenting them systematically with either a “RIGHT!” or a “WRONG!” feedback. This feedback impacted EC in the expected direction. Again, this qualifier was effective on both direct and indirect attitude measures.

That EC is sensitive to the relational meaning and the truth validation on both direct and indirect measures is highly consistent with the view that EC reflects the formations of propositions. In contrast, there is so far little evidence for unqualified CS-US linkages, failing to provide evidence for dual-learning. As we will see in Chapter Four, EC may also reflect the formation of CS-US linkages that are qualified by reciprocity. To foresee the results, the affective transfer, by default, seems to flow from the US to the CS, but not from the CSs to the USs (Mierop, Bret, Yzerbyt, Dumas, & Corneille, *in preparation*).

5.2 Stimulus-Response or Stimulus-Stimulus links

When a CS is paired with a US, it typically acquires the response elicited by the US. This could mean that what is formed in memory is a link between the mental representation of the CS and the mental representation of the US. Because the US is itself linked with a response (here, affective), this response is indirectly linked with the representation of the CS. We may therefore use the term « Stimulus-Stimulus » or S-S linkage (see Figure 2a).

Another possibility is that the mental representation of the CS is directly linked with the response elicited by the US through a « Stimulus-Response » or S-R linkage (see Figure 2b).

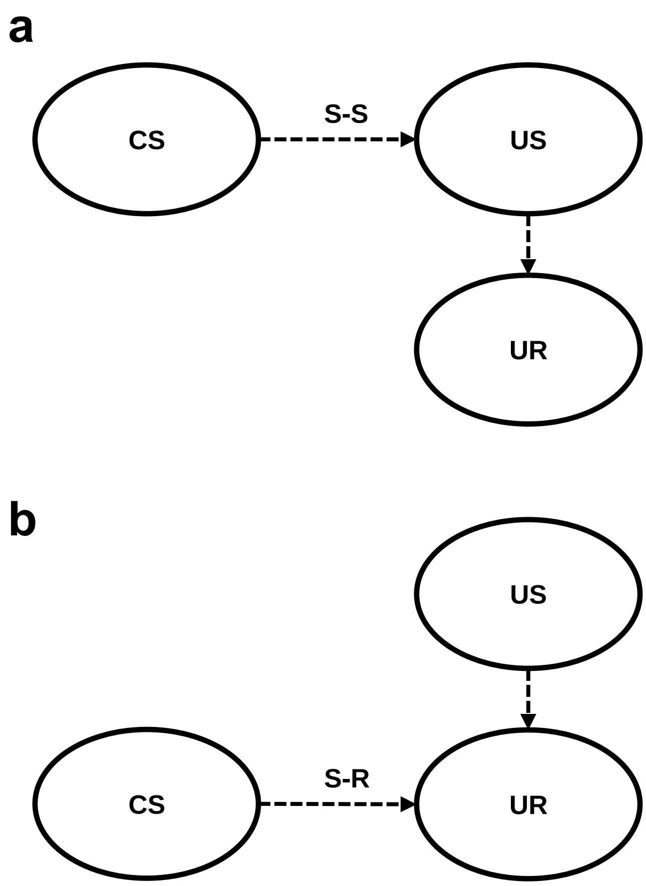


Figure 2. CS = Conditioned stimulus. US = Unconditioned stimulus. UR = Unconditioned response. S-S = Stimulus-Stimulus linkage. S-R = Stimulus-Response linkage. The mental representation of the CS is linked to the UR either directly (S-R, Figure 2b) or indirectly, through the mental representation of the US (S-S, Figure 2a).

The S-R linkage has been described as resulting from a "direct affective transfer" that can be assimilated to an associative process (Sweldens et al., 2010). It would require pairing simultaneously the CSs and the USs (Jones et al., 2009) and benefit from pairing each CS with multiple USs of a given valence (Sweldens et al., 2010). In these settings, the affective response triggered by the US can be "implicitly misattributed" to the CS (Jones et al., 2009, 2010). Importantly, this S-R account predicts that any factor increasing the ambiguity about which stimulus conveys the affective reaction should consequently increase EC. The *sine qua non* enabling condition for implicit misattribution is the simultaneous presentation of the CS and the US: the CS and the US must be processed simultaneously for the evaluative response of the latter to be misattributed to the former. Furthermore, misattribution is more likely to occur (1) with higher frequency of eye gaze shifts from one stimulus to the other, (2) with closed spatial proximity between the stimuli, (3) when the CS is more salient (i.e., bigger) than the US and (4) the US is mildly valent. Consistent with these predictions, Jones and colleagues (2009) found stronger EC effects when participants shifted more frequently their eye gaze between the CS and the US (studies 1 and 2), when the CS and US were spatially close to each other (study 3), when the CS was bigger than the US (study 4) and when the USs were mildly valent (study 5). The last observation (i.e., US intensity) is particularly important because it is to date uniquely predicted by the "implicit misattribution" account (see also Chapter Three).

What is further empirically consistent with the view that EC can rely on a S-R linkage is evidence suggesting that specific pairing procedure (i.e., simultaneous and "one-to-many" pairings) would allow "unaware" EC (Jones et al., 2009; Sweldens et al., 2010) and that EC may be insensitive to a US revaluation procedure during which the valence of the US is changed after it has been paired with a CS (Sweldens et al., 2010). This view was however challenged by experiments showing that EC is influenced by a US revaluation procedure in "conducive" settings (Frank Baeyens, Eelen, Van den Bergh, & Crombez, 1992; Walther, Gawronski, Blank, & Langer, 2009) and that it also depends on higher order processing goals (Gast & Rothermund, 2011). Furthermore, procedural conditions favoring S-R linkage does not seem to guarantee an efficient EC (Mierop et al., 2017) or a conscious-independent EC (Mierop et al., 2018; Stahl & Unkelbach, 2009) .

More generally, the S-S versus S-R distinction should not be equated to the associative versus propositional learning distinction (Corneille & Stahl, 2018; Gast & Rothermund, 2011; Sweldens, 2018). For instance, some associative accounts naturally assume that associative learning creates S-S linkages. For instance, the holistic (Levey & Martin, 1975; Martin & Levey, 1978) and the referential (Frank Baeyens et al., 1988, 1992) accounts of EC both assume that EC is due to the formation of S-S linkages that are determined by the CS-US contiguity (i.e., their absolute number of pairings). These simple associative models were followed by more complex models that assume, consistent with connectionist models (E. R. Smith, 1996; Eliot R. Smith & DeCoster, 2000), S-S linkages to be either excitatory or inhibitory (Gawronski & Bodenhausen, 2006, 2011; McConnell & Rydell, 2014).

In sum, contrary to core assumptions made in dual-process models of attitudes, there is little compelling evidence to date for either operating conditions or operating principles considered typical of EC. Yet, there are some observations that point to the existence of an associative process. Specifically, evidence for EC is found in paradigms where implicit misattribution may come into play, and some recent research suggests the possibility of uncontrolled EC; Hütter & Sweldens, 2018).

Below, I discuss four lines of research that I conducted to test further the role of associative learning in EC, by addressing its assumed operating conditions and principles.

6. Objectives and research questions

My doctoral project was aimed at contributing to the two questions discussed above, related to *when* EC is elicited, and *what* type of representation an EC procedure creates in memory. I report in this dissertation only a subset of the research lines I engaged in. Additional ones are alluded to in the General Discussion.

Chapter Two and Three relied on a Process Dissociation procedure and on MPT modeling of the data for parameters estimation. These chapters (mainly) focused on the operating conditions of EC. A distinctive feature of the studies reported in these chapters is that they combined a Process Dissociation procedure with experimental manipulations induced at encoding (i.e., attentional load and US evocativeness), as recommended by Sweldens et al. (2014). Chapters Four and Five relied on more classical behavioral measures and rather focused on the operating principles of EC. Whereas Chapter Two and Three are concerned with EC and attitude acquisition, Chapter Four and Five are rather concerned with EC and attitude change.

In Chapter Two, we tested whether an implicit learning process contributes to EC. I did so by manipulating the availability of participants' attentional resources at encoding, in a paradigm thought to be conducive to associative (i.e., unconscious and efficient) learning (Hütter & Sweldens, 2013; Jones et al., 2009; Sweldens et al., 2010).

In Chapter Three, we also tested whether an implicit learning process contributes to EC, again by relying on an "associatively-prone" paradigm. This time, we manipulated the affective evocativeness of the USs. Lowering the affective evocativeness of the US has been suggested to be one of the most crucial, yet rarely empirically addressed, facilitator of implicit EC.

In Chapter Four, we tested whether pairing CSs with USs results in reciprocal or non-reciprocal evaluative influences. Whereas an associative approach would naturally predict unqualified and therefore reciprocal influences, a propositional approach may predict qualified and therefore non-reciprocal influence because propositions can capture not only the strength but also the structure of relations between stimuli (Lagnado et al., 2007; Shanks, 2007).

In Chapter Five, we tested whether the extinction of EC is observed under high statistical power conditions, and whether EC extinction depends on the production of conditioned responses during extinction training (e.g., Rescorla, 1997). This test provides a functional comparison of EC and Classical Conditioning effects. It further tests whether the CS-US representation that is encoded in memory captures CS-US contingencies, as suggested by some associative models.

As we will now see with the empirical chapter, little evidence (if any) for associative learning was obtained across the various studies we conducted.

Chapter Two

Resource availability and explicit memory largely determine evaluative conditioning effects in a paradigm claimed to be conducive to implicit attitude acquisition

In three experiments, we investigated how preventing the explicit encoding of CS-US pairings by imposing a secondary task at learning influences evaluative conditioning (EC) effects in a paradigm claimed to be conducive to implicit EC. We additionally used a multinomial processing tree model to examine how the resource depletion manipulation affects explicit and implicit memory contributions to EC. In all experiments, the EC effect largely vanished when a secondary task was employed that severely reduced participants' explicit memory for the CS-US pairings. Furthermore, no evidence was obtained for an implicit memory contribution to EC effects. In conclusion, the present research yields evidence for explicit learning, but no support for the contribution of implicit processes to EC, in a paradigm claimed to facilitate implicit EC.

Reference

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,* 758-767. Doi: 10.1177/1948550616687093

Resource availability and explicit memory largely determine evaluative conditioning effects in a paradigm claimed to be conducive to implicit attitude acquisition

1. Introduction

Evaluative conditioning (EC) consists of an evaluative change in a conditioned stimulus (CS) following its pairing with an unconditioned stimulus (US; (De Houwer, 2007). In attitude research, EC is classically viewed as a strong case for implicit evaluative learning (Gawronski & Bodenhausen, 2006, 2009). Recent evidence, however, questions the possibility of attitude formation independent of participants' explicit encoding of CS-US pairings in memory. For instance, both explicit memory for pairings and EC effects are reduced to non-significance when imposing a secondary task at learning (e.g., (Pleyers et al., 2009). Likewise, the EC effect does not emerge when CSs are presented at near-chance identification thresholds (Stahl et al., 2016) or in parafoveal areas of the perceptual field (Dedonder et al., 2014).

Other evidence suggests, however, that specific CS-US pairing procedures may elicit implicit EC effects (Sweldens et al., 2010). Specifically, Hütter and Sweldens (2013; (Hütter et al., 2012) demonstrated implicit EC effects in a paradigm relying on simultaneous pairings involving various USs paired with a given CS. In such settings, the affect elicited by the processing of the US may be misattributed implicitly (without explicit memory for the pairings; (Jones et al., 2009) to the CS. In turn, this may produce an EC effect independently of the explicit encoding of the pairings in memory.

The present research seeks to advance our understanding of the scope of implicit EC effects by preventing an explicit encoding of the pairings in the paradigm developed by Hütter and Sweldens. This was achieved by imposing a depleting secondary task at learning. This approach conforms to recent recommendations to manipulate awareness at encoding instead of correlating the EC effect with awareness measures possibly influenced by retrieval-related effects (Gawronski & Walther, 2012).

Relying on evaluative ratings alone, however, is inconclusive for several reasons. First, the absence of residual EC effects under resource depletion conditions may be due to a lack of power. Second, the presence of residual EC effects may be attributed to the insufficiency of the secondary task (Fiedler & Hütter, 2014). Furthermore, residual EC effects under resource depletion may be attributed either to explicit or implicit processes (or both).

The multinomial processing tree (MPT) approach introduced by Hütter et al. (2012) is well-suited to examine the latter questions. This approach, based on the process dissociation procedure developed by (Jacoby, 1991), is tailored to a memory task and allows dissociating explicit (recollection of the pairings: the m-parameter) and implicit (acquired evaluations without explicit memory of the pairings: the a-parameter) memory. Hütter and Sweldens (2013) observed evidence for both a- and m-parameters using simultaneous pairings, whereas the a-parameter was not significant when using sequential pairings.

The latter evidence is consistent with the view that some CS-US pairing procedures may be conducive to implicit EC effects. Unfortunately, little evidence exists to date regarding its sensitivity to resource depletion at encoding. Consistent with Jacoby's classic demonstration (Jacoby, Toth, & Yonelinas, 1993), explicit memory, but not implicit memory, should decrease when participants' resources are taxed at encoding. The latter pattern would support the claim that MPT parameters can be interpreted as indicators of qualitatively different processes operating during learning despite being assessed after learning (Gawronski & Walther, 2012).

(Halbeisen & Walther, 2015) presented a first study that combined the MPT approach with a resource depletion manipulation. Unfortunately, their results do not inform our current research question for several reasons. First, Halbeisen and Walther (2015) compared the effect of different types of resource depletion instead of comparing the effect of taxing resources per se. That is, while different types of secondary tasks might differ in the degree of interference as shown by Halbeisen and Walther (2015), all types may reduce both explicit and implicit learning, one of these mechanisms, or none when compared to a control condition. Second, although Halbeisen and Walther (2015) used simultaneous pairings, each CS was repeatedly paired with a single US. As discussed above, this pairing procedure may prevent implicit EC because one of the two conditions for implicit EC was not met. Finally, the MPT model was based on only a small number of observations, challenging the reliability of their results.

In conclusion, a test of the role of resource depletion in an EC paradigm claimed to be conducive to an implicit EC effect is still lacking. Investigating the impact of resource depletion on EC effects in combination with the MPT approach allows addressing several questions of importance: (1) Is the EC effect reduced by resource depletion in a paradigm assumed to be conducive to implicit EC effects? (2) Does resource depletion decrease the m-parameter, supporting the notion that m reflects explicit encoding of CS-US pairings? (3) Does resource depletion fail to influence the a-parameter, supporting the view that a reflects implicit memory processes, which are known to arise independently of resource availability at encoding (Jacoby, 1991; Sweldens et al., 2010)? (4) Does any significant residual EC effect under resource depletion more likely reflect the role of implicit or explicit memory? A significant a-parameter in combination with a non-significant m-parameter would support the notion that EC may be observed in the absence of explicit encoding of the CS-US pairings in memory.

1.1 Overview of the experiments

The experiments relied on a common procedure. Participants pre-rated the CSs before they were exposed to pairings involving either negative USs (of multiple identities) or positive USs (of multiple identities). Between participants, resources at learning were either taxed or not. After the learning phase, participants completed the memory task tailored to the MPT approach, and only then the evaluative post-ratings that allowed for the analysis of EC.

Following Hütter and colleagues (2012), the memory task had participants report whether a given CS was paired with pleasant or unpleasant USs. Specifically, they were instructed to respond “pleasant” (“unpleasant”) if they remembered a CS was paired with positive (negative) USs. A distinct property of this task is that participants are requested to answer the memory task by drawing on their attitude toward the CS when they lack explicit memory of the pairing. That is, they select “pleasant” (“unpleasant”) for liked (disliked) CSs. In this inclusion condition, explicit and implicit memory lead to the same response and thus cannot be separated.

In order to dissociate explicit and implicit memory, Hütter et al. (2012) created exclusion conditions that reverse the responses based on one of the aforementioned processes so that explicit and implicit memory now lead to different responses. The exclusion instructions varied between experiments. In Experiments 1 and 2, participants reversed responses based on their evaluation of the CSs. In the absence of memory for the pairings, they were asked to respond “unpleasant” (“pleasant”) if they evaluated the CS positively (negatively). In Experiment 3, participants reversed memory-based responses, that is, they were asked to respond “unpleasant” (“pleasant”) if they remembered that the CS was paired with positive (negative) USs. In the absence of pairing memory, they were asked to simply report their evaluation of the CS.

The estimation of the m- and a-parameters involves comparing the frequencies of the “pleasant” and “unpleasant” responses between the inclusion and exclusion conditions. The processing trees for the attitude-exclusion and the memory-exclusion implementations are depicted in Figures 1 and 2, respectively. The variation of exclusion instructions between experiments established discriminant validity of the MPT parameters.

1.2 Effect size and statistical power

On the basis of the resource depletion effect size on EC effects (i.e., $\eta_p^2 = .11$) reported by Pleyers and colleagues (2009), using a two-tailed mixed ANOVA test with $\alpha = \beta = .05$, a sample of 62 participants is required to achieve high statistical power ($1 - \beta = .95$). We systematically collected larger samples in the present experiments to accommodate for potential data loss (e.g., as we excluded participants who failed the practice trials of the MPT task). More specifically, for each experiment, we decided to collect data of at least 100 participants and ran sessions up to the day this goal was reached.

The MPT model analyses were based on 1800 (Exp.1), 2256 (Exp.2), and 2760 (Exp.3) observations, that is, 24 CSs per participant. As a rule of thumb for sufficient power, (Klauer, Stahl, & Voss, 2011) suggested that not more than 10% of the expected category counts should be below five. With the present sample size, the smallest expected category counts amount to 74 (Exp.1), 74 (Exp.2), and 97 (Exp.3).

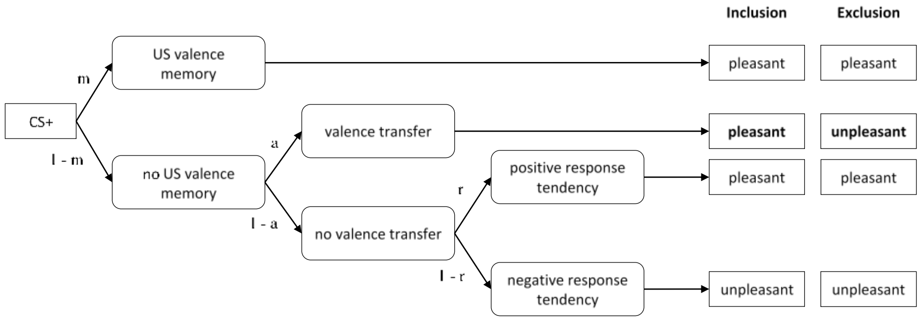


Figure 1. Processing tree model of performance in the memory task in the attitude-exclusion implementation of the MPT model. The rectangle on the left denotes the stimulus. The branches of the processing tree represent the combination of cognitive processes postulated by the model. m = probability of remembering the valence of the US; a = probability of relying on CS attitude given memory failure; r = response tendency towards “pleasant” when neither memory nor attitude are available.

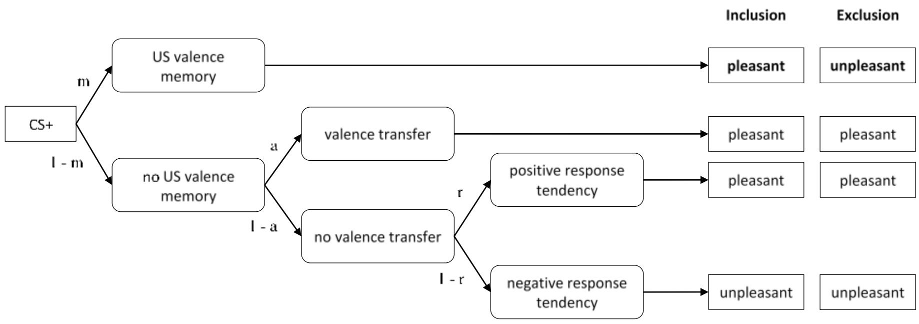


Figure 2. Processing tree model of performance in the memory task in the memory-exclusion implementation of the MPT model. The rectangle on the left denotes the stimulus. The branches of the processing tree represent the combination of cognitive processes postulated by the model. m = probability of remembering the valence of the US; a = probability of relying on CS attitude given memory failure; r = response tendency towards “pleasant” when neither memory nor attitude are available.

2. Experiment 1

Experiment 1 provides a first examination of the role of explicit encoding in a paradigm claimed to be conducive to implicit EC effects (Hütter et al., 2012, 2013; Sweldens et al., 2010). We relied on an auditory 2-back task for the resource depletion manipulation in order to reduce interferences with the visual processing of the pairings (Dedonder et al., 2010; Pleyers et al, 2009) and examined the influence of the resource manipulation on the m- and a-parameters of the MPT model.

2.1 Method

2.1.1 *Participants*

Of 100 undergraduate participants, 9 were excluded for technical reasons, 15 were excluded because they failed the memory-training task. The final sample consisted of 76 participants (M_{age} = 21.28; SD_{age} = 5.13; 70 females).

2.1.2 *Design*

A 2 (time of evaluative rating: before versus after conditioning) × 2 (US valence: positive versus negative) × 2 (Resource depletion: secondary task versus control) × 2 (memory instruction: inclusion versus exclusion) mixed design was implemented with repeated measures on the first three factors.

2.1.3 *Materials and Procedure*

The procedure combined the conditioning procedure by Hütter et al. (2012) with the distraction task used by Pleyers et al. (2009). Specifically, participants first rated 102 black-and-white portraits in a random order on a scale ranging from “unpleasant” (-100) to “pleasant” (100). Of these portraits, the 24 most neutral ones were selected as CSs. Randomly we assigned 12 CSs to positive and 12 CSs to negative US valence. During the conditioning phase, CSs were consecutively paired in a random order (though a given CS was never paired twice in a row) with eight different USs of their assigned valence, amounting to 192 pairings in total. Each picture pair was presented simultaneously for 2000ms. The inter-stimulus interval between pairs was 100ms.

The US set consisted of 50 pleasant and 50 unpleasant pictures from the IAPS (Lang, Bradley, & Cuthbert, 1997). The pleasant pictures were more positive than the unpleasant ones, $t(98) = 60.77$, $p < .001$. They did not differ in terms of arousal, $t(98) = -.02$, $p = .98$, but differed in dominance, $t(98) = 19.66$, $p < .001$. At the participant level, each USs was therefore presented up to twice.

During the conditioning phase, all participants wore headphones in which either neutral music (control condition) or randomly generated numbers from 1 to 9 (secondary task condition) were played. In the control condition, participants were informed that music would be played, whereas in the secondary task condition, participants were instructed to press the space bar each time they heard the same number twice intercalated by a different one (i.e., 2-back task, see Pleyers et al., 2009).

After the conditioning phase, the memory task was administered in the attitude-exclusion version described above (see Hütter et al., 2012, for a more extensive description of the procedure). An MPT model (Batchelder & Riefer, 1999) applied to the memory task to estimate an explicit memory parameter (m-parameter), an implicit memory parameter (a-parameter), as well as a response tendency parameter (r-parameter) as displayed in Figure 1.

Before the actual trials started, participants were presented with eight hypothetical scenarios to probe for their comprehension of the instructions. If participants made an error, instructions were repeated up to two times. If participants still made errors in the third round, they were excluded from analysis.

Finally, evaluative ratings of the 24 CSs were collected a second time. After the completion of the experiment, participants were debriefed, thanked, and dismissed. No additional measures were collected.

2.2 Results

2.2.1 *Evaluative Ratings*

CS ratings were submitted to a 2 (US valence) by 2 (time of measurement) by 2 (Resource depletion) General Linear Model (GLM), with repeated measures on the first two factors. This analysis revealed a significant EC effect (see Table 1 and Figure 3), as indicated by a significant interaction between US valence and time of measurement, $F(1,74) = 13.09$, $p < .001$, $\eta_p^2 = .15$, 90% Confidence Interval (CI) [.05, .27], observed power = .95. Importantly, the three-way interaction between US valence, time of measurement, and resource depletion was significant, $F(1,74) = 4.16$, $p < .05$, $\eta_p^2 = .05$, 90% CI [.00, .15], observed power = .52. Decomposition of this interaction by resource depletion condition revealed a significant EC effect in the control condition $F(1,41) = 14.14$, $p < .001$, $\eta_p^2 = .26$, 90% CI [.08, .41], observed power = .96, but not in the resource depletion condition $F(1,33) = 1.68$, $p = .20$, $\eta_p^2 = .05$, 90% CI [.00, .20], observed power = .24.

Table 1. Mean evaluative ratings (and standard deviations) of CSs as a function of time of measurement, US valence, and resource depletion for Experiments 1, 2, and 3.

	Pre-ratings		Post-ratings	
	CSs-	CSs+	CSs-	CSs+
Experiment 1				
Control	104.00 (6.03)	103.68 (5.86)	104.83 (14.78)	113.26 (12.69)
Secondary Task	106.38 (9.69)	105.78 (10.68)	106.01 (17.87)	107.86 (16.86)
Experiment 2				
Control	102.43 (6.32)	102.71 (6.44)	97.20 (18.26)	108.71 (14.25)
Secondary Task	101.64 (4.25)	103.01 (3.84)	98.16 (14.58)	104.56 (11.27)
Experiment 3				
Control	103.78 (6.63)	102.32 (6.38)	101.86 (6.63)	112.86 (13.47)
Secondary Task	103.78 (7.77)	103.72 (7.00)	103.58 (13.72)	105.34 (13.08)

Note. US = Unconditioned stimulus; CSs = Conditioned stimuli.

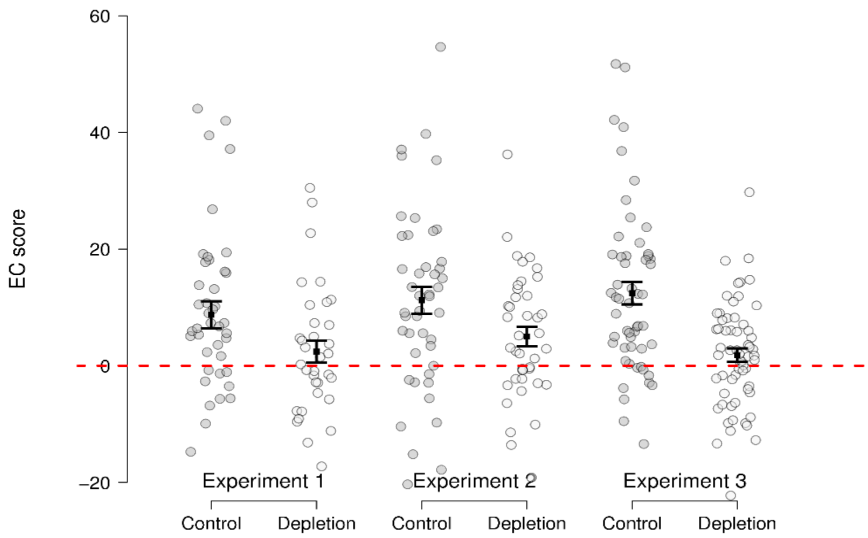


Figure 3. Mean EC effect (and 95% CIs) as a function of resource depletion (resource depletion) for Experiments 1, 2, and 3. EC score = (CSs+post - CSs+pre) + (CSs-pre - CSs-post).

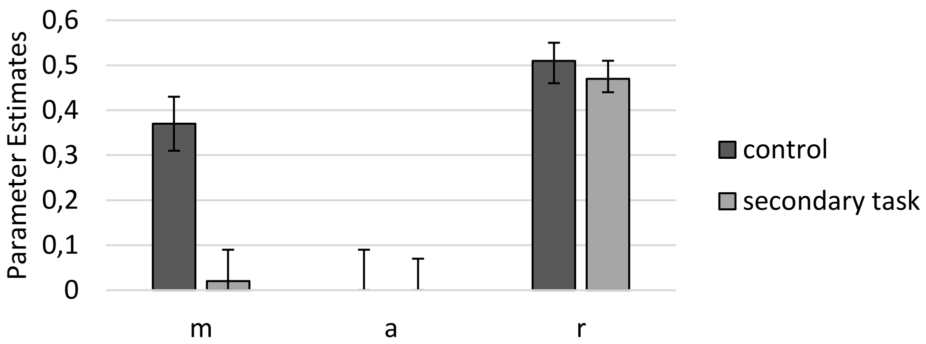
2.2.2 *MPT Model*

The initial model contained three parameters per resource depletion condition as presented in Figure 4, $G^2(2) = 13.06$, $p < .01$. We compared the parameters across resource depletion conditions by setting them equal and observing the impact on model fit. The m-parameter was significantly smaller in the secondary task condition as compared to the control condition, $\Delta G^2(1) = 56.99$, $p < .001$. In contrast, the a-parameters, $\Delta G^2(1) = 0.00$, $p = .99$, and r-parameters, $\Delta G^2(1) = 1.64$, $p = .20$, could be set equal across conditions without loss in model fit. The final model contained two m-parameters (control: m = .37, 95% CI [.31, .43], resource depletion: m = .02, 95% CI [.00, .09]), one a-parameter (a = .00, 95% CI [.00, .06]) that did not differ from zero, $\Delta G^2(1) = 0.00$, $p = .50$,[3] and one r-parameter (r = .49, 95% CI [.46, .51]) indicating no deviation from chance level (.50). However, the model had insufficient fit, $G^2(4) = 14.70$, $p < .01$, so that the present results should be interpreted with caution.

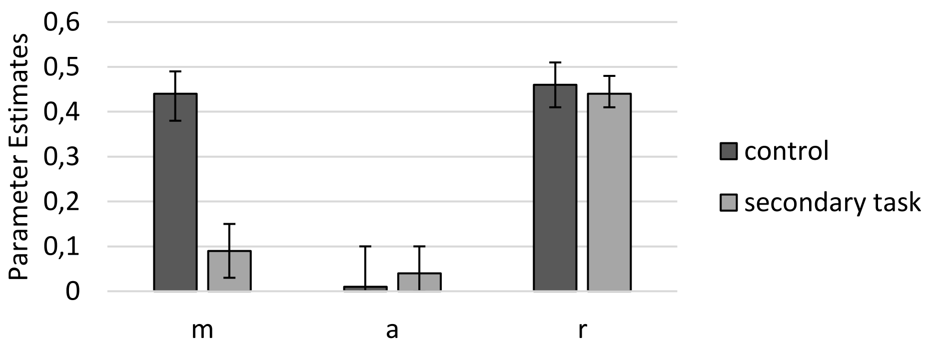
2.3 Discussion

Experiment 1 indicates that both EC and m are largely dependent on resource availability, whereas no evidence is found for implicit EC. Because the MPT model did not demonstrate good fit to the data, we realized an even closer replication of Hütter and Sweldens' (2013) original paradigm in Experiment 2.

a) Experiment 1



b) Experiment 2



c) Experiment 3

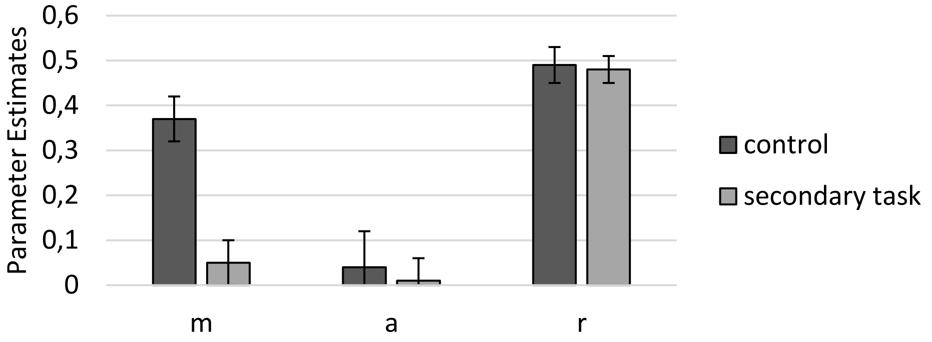


Figure 4. Estimates of the MPT model's m-, a-, and r-parameters. Error bars show the 95% confidence intervals of the parameter estimates.

3. Experiment 2

Experiment 2 implemented a slight modification in the control condition. In Experiment 1, participants wore headphones and listened to neutral music whereas they did not in Hütter and colleagues' (2012, 2013) studies. Hence, Experiment 2 removed the music (and headphones) in the control condition.

3.1 Method

3.1.1 *Participants*

We recruited 101 undergraduate students. Fifteen were excluded because they failed to answer correctly in the practice phase. The final sample consisted of 86 participants (M_{age} = 20.79; SD_{age} = 2.51; 77 females).

3.1.2 *Design, Material, and Procedure*

The design, material, and procedure of Experiment 2 were the same as in Experiment 1, with the exception that participants in the control condition did not listen to music during the conditioning phase.

3.2 Results

3.2.1 *Evaluative Ratings*

CS ratings (see Table 1 and Figure 3) were submitted to a 2 (US valence) by 2 (time of measurement) by 2 (Resource depletion) GLM, with repeated measures on the first and second factors. We again observed a strong EC effect, as shown by the significant interaction between US valence and time of measurement, $F(1,84) = 31.27$, $p < .001$, $\eta_p^2 = .27$, 90% CI [.15, .39], observed power = .99, that was moderated by the resource depletion, $F(1,84) = 4.57$, $p < .05$, $\eta_p^2 = .05$, 90% CI [.00, .14], observed power = .56. The EC effect in the resource depletion condition, $F(1,40) = 8.94$, $p < .01$, $\eta_p^2 = .18$, 90% CI [.04, .35], observed power = .83, was much smaller than in the control condition, $F(1,44) = 23.57$, $p < .001$, $\eta_p^2 = .35$, 90% CI [.16, .49], observed power = .99, although it was significant in both conditions.

3.2.2 *MPT Model*

We fitted the model with three parameters per resource depletion condition as presented in Figure 4, $G^2(2) = 2.20$, $p = .33$. The m-parameter was significantly reduced by the secondary task, $\Delta G^2(1) = 74.14$, $p < .001$. The a-parameters, $\Delta G^2(1) = 0.18$, $p = .67$, and r-parameters, $\Delta G^2(1) = 0.01$, $p = .92$, could be set equal without loss in model fit. The final model contained two m-parameters (control: m = .44, 95% CI [.39, .49], Resource depletion: m = .09, 95% CI [.03, .15]), one a-parameter (a = .03, 95% CI [.00, .08]) that did not differ from zero, $\Delta G^2(1) = 1.24$, $p = .13$, and one r-parameter (r = .45, 95% CI [.42, .48]). This model constituted a good account of the memory data, $G^2(4) = 2.71$, $p = .61$.

3.3 Discussion

Consistent with Experiment 1, resource depletion decreased the magnitude of the EC effect. This time, the MPT model fit was satisfactory, and evidence obtained in Experiment 1 replicated. The m-parameter proved highly sensitive to resource depletion, yet clearly survived it – in line with the residual EC effect in the resource depletion condition. The a-parameter was again not significant. These findings support the view that EC effects were mainly driven by explicit encoding processes.

4. Experiment 3

Experiment 3 replicates Experiment 2 with an alternative implementation of the MPT task to achieve convergent validity. Participants now were instructed to reverse their responses based on their memory instead of their evaluation in the exclusion condition of the memory task.

4.1 Method

4.1.1 *Participants*

Of 124 undergraduate participants, nine were excluded because of poor performance in the practice phase. The final sample consisted of 115 participants ($M_{age} = 19.99$; $SD_{age} = 1.32$; 105 females).

4.1.2 *Design, Material, and Procedure*

The design, material, and procedure of Experiment 3 replicated Experiment 2 with the exception that exclusion instructions asked participants to reverse the responses based on memory rather than attitudes (cf. Hütter et al., 2012). Figure 2 illustrates the predictions of the parameters.

4.2 Results

4.2.1 *Evaluative Ratings*

Evaluative ratings (see Table 1) were submitted to a 2 (US valence) by 2 (time of measurement) by 2 (resource depletion) GLM with repeated measures on the first two factors. The analysis revealed a significant EC effect, $F(1,113) = 42.28$, $p < .001$, $\eta_p^2 = .27$, 90% CI [.16, .37], observed power = .99, that was moderated by resource depletion, $F(1,113) = 23.39$, $p < .001$, $\eta_p^2 = .17$, 90% CI [.08, .27], observed power = .99. An EC effect was observed in the control condition $F(1,53) = 41.14$, $p < .001$, $\eta_p^2 = .44$, 90% CI [.26, .56], observed power = .99, but not in the secondary task condition $F(1,60) = 2.50$, $p = .12$, $\eta_p^2 = .04$, 90% CI [.00, .14], observed power = .34 (cf. Figure 3).

4.2.2 *MPT Model*

According the same analytic procedure, the model (see Figure 4) was highly consistent with the previous results, $G^2(2) = 2.09$, $p = .35$. While the m-parameter was sensitive to the resource depletion manipulation, $\Delta G^2(1) = 79.52$, $p < .001$, the a-parameters, $\Delta G^2(1) = 0.57$, $p = .45$, as well as the r-parameters, $\Delta G^2(1) = 0.11$, $p = .74$, did not differ between conditions. The final model containing two m-parameters (control: m = .37, 95% CI [.32, .42], Resource depletion: m = 0.05, 95% CI [.00, .10]), one a-parameter (a = .02, 95% CI [.00, .06]) that did not differ from zero, $\Delta G^2(1) = 0.78$, $p = .19$, and one r-parameter (r = .48, 95% CI [.46, .51]), demonstrated good fit, $G^2(4) = 2.77$, $p = .60$.

4.3 Discussion

The findings obtained in Experiments 1 and 2 generalize to a different MPT procedure. The model fit was satisfactory and conclusions turned out to be similar to those observed in Experiment 2: the m-parameter was sensitive to resource depletion and no significant a-parameter was observed. Resource depletion weakened the EC effect, again to non-significance.

5. Integrative Data Analysis

5.1 Evaluative ratings

We conducted an Integrative Data Analysis (IDA; Curran & Hussong, 2009) on the aggregated data of the three experiments. This analysis allows to increase statistical power and to test whether the effects vary across experiments. We submitted CS ratings to a 2 (US valence) by 2 (time of measurement) by 2 (Resource depletion) by 3 (Experiment) GLM with repeated measures on the first two factors.

The IDA showed that EC effect was generally strong, $F(1,271) = 79.33$, $p < .001$, $\eta_p^2 = .226$, 90% CI [.16, .29], observed power = .99, independent of Experiment, $F(2,271) = .80$, $p = .45$, $\eta_p^2 = .006$, 90% CI [.00, .02], observed power = .19. The overall effect of the resource depletion on EC was strong, $F(1,271) = 24.41$, $p < .001$, $\eta_p^2 =.08$, 90% CI [.04, .14], observed power = .99, irrespective of Experiment, $F(2,271) = .98$, $p = .38$, $\eta_p^2 = .01$, 90% CI [.00, .03], observed power = .22.

The EC effect was significant in both conditions but was much stronger in the control condition, $F(1,138) = 73.40$, $p < .001$, $\eta_p^2 = .35$, 90% CI [.24, .44], observed power = .99, than in the resource depletion condition, $F(1,133) = 12.12$, $p < .001$, $\eta_p^2 = .08$, 90% CI [.02, .16], observed power = .93.

5.2 MPT model

We conducted an integrative analysis across all 6817 observations of the three experiments. The fitted model with two m-parameters (control: m = 0.40, 95% CI [.37, .43], resource depletion: m = 0.06, 95% CI [.02, .09]), one a-parameter (a = .00, 95% CI [.00, .03]) that did not differ from zero, $\Delta G^2(1) = 0.01$, $p = .46$, and one r-parameter (r = .47, 95% CI [.46, .49]), demonstrated satisfactory fit, $G^2(20) = 29.15$, $p = .08$. Based on this analysis, the residual EC effect in the resource depletion condition is likely due to residual explicit memory for the pairings.

6. General Discussion

In the present research, we examined how depleting participants' resources at learning influences the EC effect, as well as indicators of implicit and explicit memory, in the context of a paradigm claimed to be conducive to implicit EC effects. In three experiments, despite significant variations of a common procedure (i.e., different implementations of both the MPT instructions and the control condition), we systematically reached the same conclusion: taxing participants' resources largely reduces both explicit memory for the CS-US pairings and the magnitude of EC effects. We found these effects although the encoding task and the secondary task did not overlap in sensory modality, demonstrating that these processes draw on a common resource.

We failed to obtain evidence for the contribution of implicit processes to EC, despite the use of high-powered experiments and a close replication of past procedures (Hütter & Sweldens, 2013; Hütter et al., 2012). The lack of replication of the a-parameter suggests a lack of robustness of past evidence pointing to the contribution of implicit memory to EC. As a consequence of this replication failure, an assessment of the a-parameter's sensitivity to resource depletion is not possible in the present research. However, the generally good fit of the model and the a-parameter's insensitivity to decreases in m supports the relevance of the MPT approach. That is, as proposed and empirically validated by Hütter et al. (2012, 2013), the a-parameter is not affected by a reduction in the m-parameter, suggesting that the two parameters represent qualitatively different processes rather than processes that differ in quantitative terms (e.g., the strength of the memory signal).

All three experiments show that EC is sensitive to opportunities of explicit encoding, with larger EC effects in the control condition than in the resource depletion condition. However, the absolute levels of EC observed in the evaluative ratings in these conditions are difficult to interpret. As there is no absolute scale for depletion (Fiedler & Hütter, 2014), one cannot know whether the two-back task used in the present research fully depleted participants' resources. Thus, the small (and often non-significant) EC effects under resource depletion could be due to processes operating implicitly or residual resources that were not taxed by the secondary task. Results from the MPT model support the latter interpretation.

The present research relied on direct evaluative measures (i.e., evaluative ratings). One could argue that EC effects assessed with indirect measures (Fazio, Jackson, Dunton, & Williams, 1995; Payne, Cheng, Govorun, & Stewart, 2005) may have proven less sensitive to resource depletion. However, several theoretical considerations led us to focus on direct evaluations in the present research. First, the effect size of EC effects is considerably larger on direct than on indirect measures (Hofmann et al., 2010). Finding no evidence for EC effects on an indirect measure may thus have been attributed to a lack of power. Second, evidence is accumulating that questions the existence of qualitative differences between direct and indirect measures. Effects consistent with explicit encoding were obtained on indirect measures (e.g., De Houwer, 2006; Pleyers et al., 2007; Stahl, Unkelbach, & Corneille, 2009) effects consistent with implicit learning were obtained on direct measures (Hütter et al., 2012; Jones et al., 2009; Olson & Fazio, 2001). Thus, the nature of the measure is not necessarily informative with regard to the type of learning underlying evaluation. The MPT measure relies on self-reports that allow for flexible responding in line with instructions and allows for keeping the task constant. Nevertheless, it is an open research question whether indirect measures of evaluation would show the same pattern of results.

The contribution of this research is manifold. In general, the experiments reported here are consistent with recent invitations to address the role of contingency awareness in EC through experimental manipulations at the encoding stage (Gawronski &Walther, 2012) and to combine experimental approaches with MPT modeling (Sweldens et al., 2014). Additionally, the present research allows reaching a number of more specific innovative conclusions. First, it reveals that the detrimental impact of resource depletion extends to a paradigm thought to be conducive to implicit learning processes. Second, it shows that explicit memory for the CS-US pairings plays a large role in this paradigm. Third, it further validates the view that the m-parameter reflects encoding-related processes (as m was sensitive to an encoding-related manipulation) and that m can be varied independently of the a-parameter. Fourth, it suggests that the estimation of the a-parameter in the present paradigm is less robust than initially thought.

Nevertheless, the current experiments should not be interpreted as revealing the general absence of implicit processing in EC. First, the question of whether an implicit process operates in the current paradigm remains open. In two unrelated studies using the same conditioning paradigm, we recently observed a significant a-parameter. We may therefore conclude that the a-parameter is obtained in the current paradigm, although it seems not to be quite robust. Furthermore, another CS-US pairing procedure may be more suitable for obtaining implicit EC effects. However, such evidence is currently scarce. For instance, subliminal EC studies generally suffered from design-related issues, lacked sensitive awareness tests, and operated under exposure conditions that have been shown to allow conscious identification (Lovibond & Shanks, 2002; Pleyers et al., 2007; Stahl et al., 2016; Sweldens et al., 2014). When tightly controlled CS exposure conditions and sensitive awareness checks are used, meeting methodological requirements currently advised in cognitive psychological research (Atas, Vermeiren, & Cleeremans, 2013; Pratte & Rouder, 2009; Shanks & St. John, 1994), objective identification appears to be a necessary yet insufficient condition for an EC effect to emerge (Stahl et al., 2016). As to subliminally presented affective stimuli (i.e., USs), they have recently been shown to elicit no affective response (Lähteenmäki, Hyönä, Koivisto, & Nummenmaa, 2015).

As a final comment, evidence supporting the possibility of an implicit EC effect stems from incidental learning studies by Fazio and colleagues (Jones et al., 2009; Olson & Fazio, 2001). Their paradigm involved a distracting vigilance task and an even larger number of different USs paired with the CSs. In addition, the USs used in these studies likely involved USs of lower intensity. Assuming the latter effect is robust and survives stringent awareness tests, the scope of implicit EC effects observed in the vigilance paradigm would however exclude from the realm of implicit EC effects: backward and forward CS-US pairings, non-incidental exposures, procedures involving USs of either too weak or too high intensity, and of either too large or too small perceptual salience. Such a small scope of implicit EC would severely limit the practical and theoretical interest of dual-process models of attitude acquisition and perhaps of implicit learning in general (Shanks, 2005).

Chapter Three

Does attitude acquisition in evaluative conditioning without explicit CS-US memory reflect implicit misattribution of affect?

Research that dissociates different types of processes within a given task using a processing tree approach suggests that attitudes may be acquired through evaluative conditioning in the absence of explicit encoding of CS-US pairings in memory. This research distinguishes explicit memory for the CS-US pairings from CS-liking acquired without encoding of CS-US pairs in explicit memory. It has been suggested that the latter effect may be due to an implicit misattribution process that is assumed to operate when US evocativeness is low. In the present research, the latter assumption was supported neither by two high-powered experiments nor by complementary meta-analytic evidence, whereas evocativeness exerted an influence on explicit memory. This pattern of findings is inconsistent with the view that CS-liking acquired without encoding of CS-US pairs in explicit memory reflects an implicit misattribution process at learning. Hence, the underlying learning process is awaiting further empirical scrutiny.

Reference

Mierop, A., Hütter, M., Stahl, C., & Corneille, O. (2018). Does attitude acquisition in evaluative conditioning without explicit CS-US memory reflect implicit misattribution of affect? *Cognition and Emotion.* doi: 10.1080/02699931.2018.1435505

Does attitude acquisition in evaluative conditioning without explicit CS-US memory reflect implicit misattribution of affect?

1. Introduction

When a neutral stimulus (conditioned stimulus; CS) is paired with a positive or negative stimulus (unconditioned stimulus; US), the CS may acquire the US valence. This attitude formation phenomenon, coined "evaluative conditioning" (Martin & Levey, 1978), has been studied extensively (e.g., Hofmann et al., 2010). An intensely debated question is whether this simple attitude learning mechanism can occur independently of the encoding of CS-US pairings in explicit memory.

Experimental and correlational research has typically provided conflicting answers to that question (for a recent review, see Sweldens et al., 2014; Corneille & Stahl, 2017). Experimental studies that prevented the explicit encoding of CS-US pairings in memory generally found weak evidence for memory-independent EC. This conclusion was supported in experiments that failed to obtain a significant EC effect when participants' cognitive resources were taxed at learning (e.g., Davies et al., 2012; Dedonder et al., 2010; Mierop et al., 2017; Pleyers et al., 2009; Kattner, 2012) or when participants were exposed to briefly presented (e.g., Hofmann et al., 2010; Stahl et al., 2016; see however Greenwald & De Houwer, 2017), parafoveally presented (Dedonder et al., 2013), or visually suppressed (Högden et al., 2018) stimuli.

Turning to correlational evidence, studies that relied on item-based analyses of CS-US memory generally found EC effects only for CSs that were successfully related to their associated US identity (e.g., Pleyers et al., 2007) or US valence (e.g., Stahl et al., 2009) in a later memory task. This finding was obtained in various sensory modalities (Wardle, Mitchell, & Lovibond, 2007) and was supported using psychophysiological measures (Dawson, Rissling, Schell, & Wilcox, 2007). In contrast, however, studies relying on a more refined processing tree approach using multinomial processing tree (MPT) modeling, found evidence for EC in the absence of explicit memory for the pairings (Hütter et al., 2012). The present research is the result of an adversarial collaboration that examined the role of implicit affective misattribution as a driver of the latter implicit learning evidence, by manipulating one of its moderators: US evocativeness.

The contribution of implicit affect misattribution to EC without CS-US memory.

Hütter and colleagues (2012) noted that when participants report the valence of a US paired with a CS, both valence memory ("I remember this image was paired with a positive image") and inferences ("I like this image, so it must have been paired with a positive image") may contribute to US valence responses. A process dissociation approach (Jacoby,

1991) allows disentangling the contribution of these independent processes. This is achieved by contrasting an “inclusion” to an “exclusion” instruction condition, where these two processes produce convergent or divergent responses, respectively. As explained below (see Methods section), data generated under these conditions may then be modeled to yield separate estimates for memory-dependent (recollection of the CS-US pairings: the m-parameter) and memory-independent (acquired evaluations without explicit memory of the pairings: the a-parameter) effects in EC.

Hütter and colleagues (2012) showed that the estimate of the explicit memory component, parameter m, was larger than zero, suggesting that participants successfully encode a substantial proportion of the CS-US pairings in explicit memory. More importantly, they also found that the estimate of memory-independent EC effects, parameter a, was larger than zero, suggesting that participants may acquire CS liking consistent with US valence in the absence of explicit CS-US memory. Of critical interest, Hütter and Sweldens (2013) further predicted and found the EC effect on the a-parameter to be obtained in simultaneous, but not in sequential CS-US pairing settings.

These authors explained that the latter finding is consistent with the operation of an implicit affective misattribution (IM) process (Jones et al., 2009, 2010). That is, given simultaneous CS-US pairings, participants may be confused about the source of their affect at learning: they would implicitly and mistakenly attribute their affective response to the CS instead of the US. As a result of this IM process, the CS may acquire the US valence in the absence of encoding of the CS-US pairing in explicit memory. Presumably, this would result in a memory-independent EC effect that is captured by the a-parameter.

While temporal contiguity is considered a particularly relevant and critical factor to IM by Jones et al. (2009), simultaneous versus sequential pairings, however, are also potentially relevant for other processes contributing to EC. That is, the temporal proximity of the CS and US might foster a multitude of attitude acquisition processes. In particular, the a-parameter might also represent associative learning of mere referential relations (Baeyens et al., 1992) or the formation of a holistic representation (Martin & Levey, 1978). Hence, it is possible that other such processes conceptualized as driving EC effects in the absence of explicit memory also profit from simultaneous over sequential presentations. The reasoning that a multitude of different processes may be affected by temporal contiguity also receives support from variation of the m-parameter in Hütter and Sweldens’ (2013) study, which was larger when CS-US pairs were presented simultaneously rather than sequentially. As a consequence, a temporal contiguity manipulation does not allow for a conclusive evaluation of the discriminant validity of the a-parameter (Campbell & Fiske, 1959).

Yet, it may still be the case that the EC effect reflected in the a-parameter partly reflects IM in a simultaneous CS-US pairing paradigm. This possibility can be addressed by manipulating other factors that are relevant to IM and that are uniquely postulated by that account. In the present research, we chose to manipulate US evocativeness, which is considered another critical determinant of IM. US evocativeness was examined as a moderator of IM in an experiment conducted by Jones and colleagues' (2009). These authors paired CSs with USs that were pretested to be rated as more intense on an evaluative dimension and categorized faster as 'positive' or 'negative' in a dichotomous categorization task (strongly evocative USs) versus to be rated as relatively milder and categorized slower as 'positive' or 'negative' (weakly evocative USs).

In that experiment, low evocativeness decreased pairing memory, yet increased EC effects. Presumably, low evocativeness enhanced source confusability, thereby boosting the operation of IM in EC. Importantly, this proposed pattern is unique to the IM account. Even if one would allow for an influence of US evocativeness on memory-independent processes as conceptualized in the referential (Baeyens et al., 1992) and holistic accounts (Martin & Levey, 1978), one would expect that lower evocativeness would decrease EC effects as learning effects are assumed associative in nature and thus reflective of the properties of the US.

In the present adversarial collaboration, we examined the latter hypotheses in two experiments that relied on the same CS material, conditioning procedure, and memory task as in Hütter and colleagues (2012). US evocativeness was manipulated between-participants in Experiment 1 and within-participants in Experiment 2. If low evocativeness facilitates IM, and if the a-parameter reflects IM, then it should be the case that a greater a-parameter is obtained for CSs simultaneously paired with weakly as compared to strongly evocative USs. In contrast, obtaining a lower or an equivalent a-parameter in the low evocativeness condition as compared to the high evocativeness condition would challenge the claim that the memory-independent contribution to EC reflects the operation of an IM process. As to the m-parameter, its estimate may be found to be lower for CSs paired with weakly than strongly evocative USs (Jones et al., 2009).

2. Experiment 1

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. The data from the experiment as well as from the pilot study for the selection of the USs are publicly available on Open Science Framework (osf.io/6bjvy).

2.1 Method

2.1.1 *Participants*

The sample size of this experiment was estimated on the basis of Hütter and Sweldens (2013), who relied on 77 participants. We collected six more participants to accommodate potential data loss (participants are excluded if they fail the training stage). Among the 83 initial participants, seven were excluded because they failed the practice trials preceding the memory task three times in a row. The final sample was therefore composed of 76 participants (M_{age} = 21.21; SD_{age} = 2.24; 62 female). The number of observations collected respects the “rule of thumb” for sufficient power suggested by Klauer, Stahl, and Voss (2011), who recommended that not more than ten percent of the expected category counts should be below five.

2.1.2 *Design*

A 2 (time of evaluative rating: before versus after conditioning) × 2 (US valence: positive versus negative) × 2 (US evocativeness: low versus high) × 2 (memory instruction: inclusion versus exclusion) mixed design was implemented with repeated measures on the first two factors.

2.1.3 *Procedure*

The procedure was identical to Hütter and colleagues (2012; Experiment 2a), but for the USs used. Specifically, the USs were 200 affective pictures (all coming from the International Affective Picture System, IAPS, Lang, Bradley, & Cuthbert, 1999, see Appendix) that were selected on the basis of a pilot study. These 200 USs were composed of 50 USs of each of four (valence × evocativeness) categories. Consistent with the evocativeness criteria used by Jones and colleagues (2009, Experiment 5), we measured both the self-reported liking and the time needed to categorize each US as “positive” or “negative”. The stimuli of low and high evocativeness (see Table 1) differed both in (absolute value) ratings, $F(1,198) = 6.34$, $p = .013$, $\eta_p^2=.03$, and in the amount of time needed to categorize them as positive or negative, $F(1,198) = 244.59$, $p < .001$, $\eta_p^2=.55$. They, however, differed neither in categorization accuracy nor in arousal (Fs < 1). The effect size of the evocativeness manipulation (at least for the direct evaluative ratings) was modest, but the mean values closely followed the mean ratings of the stimuli selected by Jones and colleagues (2009).

The CSs were 24 black-and-white portraits selected on a participant level. They were the most neutral portraits for each participant among 102 that were pre-rated in a random order on an evaluative scale ranging from “very unpleasant” (-100) to “very pleasant” (100).

The CSs were randomly assigned to USs of positive (12 CS+) or negative valence (12 CSs-). A given CS was paired with eight different USs of their assigned valence during the conditioning phase, amounting to 192 pairings in total. Each picture pair was presented simultaneously for 2000 ms with an inter-stimulus interval between pairs of 100 ms.

After the conditioning phase, participants performed the memory task. They were randomly assigned to the inclusion or the exclusion condition. The task of the participants was, for each CS, to report whether it had been paired with pleasant or unpleasant pictures by selecting one of the two answer boxes (“pleasant” and “unpleasant”). Participants in the inclusion condition were instructed to respond in accordance with their memory. In the exclusion condition, participants were instructed to reverse their responses according to their memory of the pairings (e.g., by selecting the “pleasant” answer if a CS was paired with unpleasant USs). Importantly, participants in both conditions were also instructed to report their attitude toward the portrait (without reversing their response) when they did not remember the pairings. Participants were never instructed to guess.

Before the actual memory task started, participants were trained with eight hypothetical scenarios to probe for the comprehension of the instructions. If participants made errors, instructions were repeated up to two times. Participants who still made errors in the final repetition were excluded from the analysis in accordance with Hütter at al. (2012) and Hütter and Sweldens (2013). Finally, evaluative ratings of the 24 CS faces were obtained a second time. After the completion of the experiment, participants were thanked and dismissed.

Table 1. Mean evaluative ratings, categorization time, categorization accuracy, and arousal ratings for the positive and negative USs of low and high evocativeness.

	Low evocative USs		High evocative USs	
	Negative	Positive	Negative	Positive
Experiment1				
Ratings (1-9)	3.4 (.8)	6.5 (.6)	3.1 (.6)	6.9 (.8)
RT (ms)	986.5 (137.2)	1053.6(141.4)	786.5 (51.0)	790.9 (22.1)
Accuracy (0-1)	.8 (.2)	.8 (.1)	.9 (.1)	.9 (.1)
Arousal (1-9)	5.2 (.9)	4.5 (.9)	5.3 (.9)	4.5 (.9)
Experiment2				
Ratings (1-9)	4.0 (.5)	5.9 (.3)	2.5 (.2)	7.4 (.3)
RT (ms)	1034.5(165.9)	1109.9 (177.0)	787.1 (48.2)	794.0 (20.9)
Accuracy (0-1)	.7 (.2)	.8 (.1)	.9 (.1)	1.0 (.1)
Arousal (1-9)	4.9 (1.0)	4.2 (.9)	5.6 (.8)	4.7 (1.0)

Note. USs = Unconditioned stimuli; CSs = Conditioned stimuli. RT = response time (i.e., time in milliseconds to categorize a stimulus as “positive” or “negative”).

2.2 Results

2.2.1 *Evaluative Ratings*

Evaluative ratings of the CSs were submitted to a 2 (US valence) by 2 (time of measurement) by 2 (US evocativeness) ANOVA, with repeated measures on the first and second factors. This analysis revealed an EC effect, as indicated by a significant interaction between US valence and time of measurement, $F(1, 74) = 81.16$, $p < .001$, $\eta_p^2 = .52$.

Positively paired CSs were rated more positively than negatively paired CSs in the postratings, $F(1, 74) = 86.66$, $p < .001$, $\eta_p^2=.54$, but not in the preratings, $F(1, 74) = .09$, $p = .76$, $\eta_p^2 = .00$. The three-way interaction between US valence, time of measurement, and evocativeness condition was not significant, $F(1, 74) = .06$, $p = .81$, $\eta_p^2 = .00$. The full pattern of means is reported in Table 2.

2.2.2 *Memory task*

We modeled the frequency data of the memory task (see Table 3) using the HMMtree software (Stahl & Klauer, 2007). We adapted the MPT model proposed by Hütter et al. (2012) by expressing the level of memory-independent EC in the high-evocativeness condition as a proportion f of the memory-independent EC in the low-evocativeness condition (i.e., $a_{\text{high-evocativeness}} = f*a_{\text{low\_evocativeness}}$). If high US evocativeness indeed harms memory-independent EC as predicted by the IM account, parameter f should be significantly smaller than 1. The parameters of interest were therefore a (reflecting the level of memory-independent EC) and f (reflecting the effects of the evocativeness manipulation on the a-parameter). In addition, for each evocativeness condition, the model estimated one m-parameter (reflecting explicit memory) for each US valence (positive, negative) to allow for valence asymmetries, and one r-parameter (reflecting the tendency to guess “pleasant”).

Restricting the a-parameter to zero tests the hypothesis that no memory-independent EC effect is present. The hypothesis is rejected if this restriction leads to significant reduction of the model’s goodness of fit. The hypothesis that memory-independent EC effects were absent must be rejected based on the fact that a could not be fixed at zero, $\Delta G^2(1) = 7.01$, $p = .004$.

Restricting parameter f to a constant value of 1 tests the hypothesis that memory-independent EC is not affected by US evocativeness (i.e., is of the same magnitude with low- and high-evocative USs). The hypothesis is rejected if this restriction leads to significant reduction of the model’s goodness of fit. The parameter f = .71, 95% CI [.00, 1] could be fixed at 1 without reducing goodness of fit, $\Delta G^2(1) = 0.20$, $p = .33$. The null hypothesis that US evocativeness does not affect memory-independent EC could be retained ($a_{low} = .11$, 95% CI [.00, .22]; $a_{high} = .08$, 95% CI [.00, .19]).

Retaining the null hypothesis raises the issue of power and sensitivity. We conducted power analyses using *multiTree* (Moshagen, 2010) and found that the memory-independent EC effect (i.e., $a > 0$) had an effect size of $w = .06$ (Cohen, 1988) , and given our sample sizes ($N = 76$) and the observed level of explicit memory, the present studies were sufficiently sensitive to detect effects of this magnitude with alpha = .05 and power = .75. In contrast, the effect of US evocativeness on memory-independent EC was much smaller ($w = .01$). To detect such a very small effects in the present design, we would have needed 71,154 observations or 2,965 participants. Thus, we cannot exclude the possibility of very small effects of US evocativeness on EC. In the present study, however, that was sufficiently powered to detect small effects, US evocativeness failed to reduce the MPT estimate of memory-independent EC substantially.

Next, we assessed the moderation of the m-parameters by evocativeness condition. The m-parameter for positively paired CSs was not influenced by evocativeness, such that memory was comparable between the low-, m_{low_pos} = .37, 95% CI [.29, .46], and high-evocativeness conditions, m_{high_pos} = .46, 95% CI [.37, .54], $\Delta G^2(1) = 1.83$, $p = .18$. Evocativeness also did not influence memory for negative pairings, $\Delta G^2(1) = 0.11$, $p = .18$. The m-parameter for CSs- in the low-evocativeness condition amounted to m_{low_neg} = .50, 95% CI [.42, .58], and m_{high_neg} = .48, 95% CI [.40, .56] in the high-evocativeness condition. When equated across evocativeness conditions, the resulting m-parameters for positive, m_{pos} = .41, 95% CI [.36, .47], and negative pairings, m_{neg} = .49, 95% CI [.44, .55], demonstrated a marginal difference, $\Delta G^2(1) = 3.52$, p = .06.

Finally, we explored effects of evocativeness on guessing. The r-parameters differed significantly between evocativeness conditions, $\Delta G^2(1) = 4.07$, $p = .04$. While the r-parameter in the low-evocativeness condition did not differ significantly from .50, r_{low} = .52, 95% CI [.46, .58], $\Delta G^2(1) = 0.00$, $p = .99$, for the high-evocativeness condition the r-parameter indicated a negative response tendency, r_{high} = .43, 95% CI [.37, .49], $\Delta G^2(1)$ = 5.09, $p = .02$.

Table 2. Mean evaluative ratings (and standard deviations) of CSs as a function of time of measurement, US valence, and US evocativeness.

	Preratings		Postratings	
	CSs-	CSs+	CSs-	CSs+
Experiment 1				
Mild USs	103.08(6.72	102.30(6.15	95.05(15.04)	108.82(13.28
Intense USs	102.49(3.94	103.54(4.43	98.20(10.70)	114.60(10.49
Experiment 2				
Mild USs	103.61(8.82	104.37(8.76	101.64(18.21	112.98(19.15
Intense USs	104.35(8.76	103.50(8.51	100.17(19.15	113.40(16.72

Note. USs = unconditioned stimuli; CSs = conditioned stimuli.

Table 3. Observed frequencies of “pleasant” (+) and “unpleasant” (-) responses under inclusion and exclusion instructions in the different conditions of Experiments 1 and 2.

	Inclusion		Exclusion	
	+	-	+	-
Experiment 1				
Low evocativeness				
CSs+	181	83	60	132
CSs-	50	214	133	59
High evocativeness				
CSs+	164	52	73	167
CSs-	53	163	175	65
Experiment 2				
Low evocativeness				
CSs+	342	162	180	240
CSs-	157	347	279	141
High evocativeness				
CSs+	357	147	146	274
CSs-	128	376	321	99

Note. CSs = conditioned stimuli.

2.3 Discussion

To summarize the results of Experiment 1, the memory-independent parameter (a) was not dependent on the evocativeness condition. However, as the null-effects on both the explicit memory-parameter (m) and the evaluative ratings suggest, our manipulation of evocativeness might not have been strong enough in the present experiment.

3. Experiment 2

The fact that we did not obtain effects of US evocativeness in Experiment 1 may be due to a lack of statistical power and/or a lack of strength of the evocativeness manipulation. To address this possibility, Experiment 2 increased statistical power by using a within-participants design and doubling the sample size, and strengthened the evocativeness manipulation.

Experiment 1 used 200 USs with 50 USs in each group that had relatively large variability in their evocativeness. Furthermore, the contrast between mild and intense USs was perhaps too weak because of the between-participants evocativeness manipulation. It is also possible that our set of ‘weakly evocative’ USs were in fact still too high in absolute evocativeness. For these reasons, we conducted a second experiment that made use of more differentiated US sets and manipulated US evocativeness within-participants.

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. The data from the experiment as well as from the pilot study for the selection of the USs are publicly available on Open Science Framework (osf.io/6bjvy).

3.1 Method

3.1.1 *Participants*

One hundred and seventy-eight subjects participated in this experiment. Twenty-four were excluded because they failed the memory-training task three times in a row. The final sample was therefore composed of 154 participants (M_{age} = 20.64; SD_{age} = 2.94; 139 female).

3.1.2 *Design and Procedure*

Design and procedure were identical to the first experiment with the only exception that the US evocativeness manipulation was performed within-participants and that only the mildest half of the low-evocativeness and the strongest half of the high-evocativeness USs were kept in the EC procedure. Hence, each participant was exposed to the same number of CS-US pairs as in the first experiment (i.e., 192 pairings).

We selected 25 USs per category (valence × evocativeness) resulting in 100 USs (all coming from the International Affective Picture System, IAPS, Lang et al., 1999, see Appendix). The stimuli of low and high evocativeness differed both in (absolute value) ratings, $F(1,98) = 83.53$, $p < .001$, $\eta_p^2=.46$, and in the amount of time needed to categorize them as positive or negative, $F(1,98) = 125.36$, $p < .001$, $\eta_p^2 = .56$. They differed neither in categorization accuracy nor in arousal (Fs < 1). As compared to Experiment 1, the stimuli of low and high evocativeness now differed strongly in valence ($\eta_p^2 = .46$), while the effect on the categorization measure remained strong ($\eta_p^2 = .56$).

A 2 (time of evaluative rating: before versus after conditioning) × 2 (US valence: positive versus negative) × 2 (US evocativeness: low versus high) × 2 (memory instruction: inclusion versus exclusion) mixed design was implemented with repeated measures on the first three factors.

3.2 Results

3.2.1 *Evaluative Ratings*

Evaluative ratings of the CSs were submitted to a 2 (US valence) by 2 (time of measurement) by 2 (US evocativeness) repeated-measures ANOVA. This analysis revealed a significant EC effect, as shown by a significant interaction between US valence and time of measurement, $F(1, 153) = 86.68$, $p < .001$, $\eta_p^2 = .36$. The three-way interaction between US valence, time of measurement, and evocativeness was statistically significant, $F(1, 153) = 5.16$, $p = .025$, $\eta_p^2 = .03$. However, the effect size of the interaction was small and the analyses on the level of the two evocativeness conditions demonstrated significant EC effects which were comparable in size in both the high evocativeness condition, $F(1, 153) = 70.58$, $p < .001$, $\eta_p^2 = .32$, and the low evocativeness condition, $F(1,153) = 59.25$, $p < .001$, $\eta_p^2 = .28$. The full pattern of means is reported in Table 2.

3.2.2 *Memory Task*

We used the same adapted MPT model as in Experiment 1 to model the frequency data obtained (see Table 3). As in the last experiment, the parameter a could not be fixed at zero, $\Delta G^2(1) = 8.39$, $p = .002$. Thus, the hypothesis that memory-independent EC was absent must be rejected. To assess whether memory-independent EC was affected by US evocativeness, we tested whether the parameter f (f = .27, 95% CI [0, 1]), could be fixed at 1 without reducing goodness of fit. This was indeed the case, $\Delta G^2(1) = 1.74$, $p = .19$, so that the null hypothesis stating that memory-independent EC is not affected by US evocativeness could be retained ($a_{low} = .09$, 95% CI [.03, .15]; $a_{high} = .03$, 95% CI [.00, .10]).

The memory-independent EC parameter had an effect size of w = .05 (i.e., a > 0), and given our sample size (N = 154) and the level of explicit memory, the present study was sufficiently sensitive to detect effects of this magnitude with alpha = .05 and power = .83. In contrast, the effect of US evocativeness on memory-independent EC was much smaller, w = .02. Detecting such a small effect in the present design with alpha = .05 and power = .80 would have required 16,630 observations, which is equivalent to 693 participants. Thus, again we cannot exclude the possibility of very small effects of US evocativeness on memory-independent EC.

However, in this study, we obtained evidence that the lack of an effect on the a-parameter cannot be explained by a weak or unsuccessful manipulation of US evocativeness as its effectiveness was ensured by careful pretesting. Moreover, the evocativeness manipulation affected explicit memory for US valence (i.e., the m-parameter). For positively paired CSs, the m-parameter was larger for the high-evocativeness condition, $m_{high\_pos} = .36$, 95% CI [.30, .42], than the low-evocativeness condition, $m_{low\_pos} = .25$, 95% CI [.19, .31], $\Delta G^2(1) = 6.22$, $p = .01$. The estimates for the low-, $m_{low\_neg} = .35$, 95% CI [.29, .41], and high-evocativeness conditions, $m_{high\_neg} = .51$, 95% CI [.03, .15], also differed for negatively paired CSs, $\Delta G^2(1) = 14.01$, $p = .0002$. Within both the low-, $\Delta G^2(1) = 5.34$, $p = .02$, and high-evocativeness conditions, $\Delta G^2(1) = 12.69$, $p = .0004$, memory parameters were larger for CSs- than for CSs+. In Experiment 2, the r-parameters for the low-, r_{low} = .53, 95% CI [.49, .56], and high-evocativeness conditions, r_{high} = .53, 95% CI [.49, .57], were equal in size, $\Delta G^2(1) = 0.01$, $p = .92$.

3.3 Discussion

In this second experiment, we strengthened the US evocativeness manipulation by selecting the most differentiated USs from Experiment 1 and by relying on a within-participant manipulation. We observed that the evocativeness manipulation affected explicit memory for the CS-US contiguity (the m-parameter): CS-US pairs with intensely evocative USs were better remembered than CS-US pairs with mildly evocative USs. Finally, we again observed that although the a-parameter was overall significantly different from zero, it was not affected by the evocativeness manipulation.

Although the null hypothesis (that there is no difference in the magnitude of the a-parameter due to the evocativeness manipulation) was not rejected, at this point we cannot conclude it has to be accepted. In the following section, we report a meta-analysis on the two experiments reported here as well as previously published experiments using a similar EC procedure that reported an estimate of the a-parameter. We did so to determine the statistical heterogeneity (see Higgins & Thompson, 2002) between experiments and so to examine whether the variation between the a-parameters obtained in the experiments reported here is above the variation that can be expected from sampling error.

4. Meta-analysis

The present analysis includes all published experiments in which an estimate of the a-parameter was obtained and in which the material as well as the EC procedure was comparable. The criteria for the material were: (i) black-and-white human faces as CSs, and (ii) IAPS pictures as USs. The criteria for the EC procedure were: (i) simultaneous pairing of the CS-US pairs, and (ii) CSs paired with USs that shared the valence but differed in identity (i.e., “one-to-many” pairings). Eight published data sets met these criteria. Three data sets were from the original paper by Hütter and colleagues (2012; Experiments 2a, 2b and 3), three were comprised by the control conditions from Mierop and colleagues (2017), one consisted of the “simultaneous pairings” condition from Hütter and Sweldens (2013), and one consisted of the “experienced pairings” condition from Hütter and De Houwer (2017).

We conducted a meta-analysis (see Figure 1) on the estimates of the a-parameters and their precision (i.e., confidence intervals) using RStudio (RStudio Team, 2015) and the ‘meta’ package (Schwarzer, 2007). By doing so, we were able to approximate the overall magnitude of the a-parameter and, perhaps more important for our current purpose, the degree of heterogeneity between experiments. If the estimates of the a-parameters obtained in our experiments fall into the distribution of the a-parameters from the referent experiments and do not increase the heterogeneity, we might conclude that different evocativeness conditions should not be thought of as influencing the size of the a-parameter. By contrast, if they do increase the heterogeneity between experiments, then differences should not be seen as resulting from chance. The R script for the meta-analysis is publicly available via the Open Science Framework (osf.io/6bjvy).

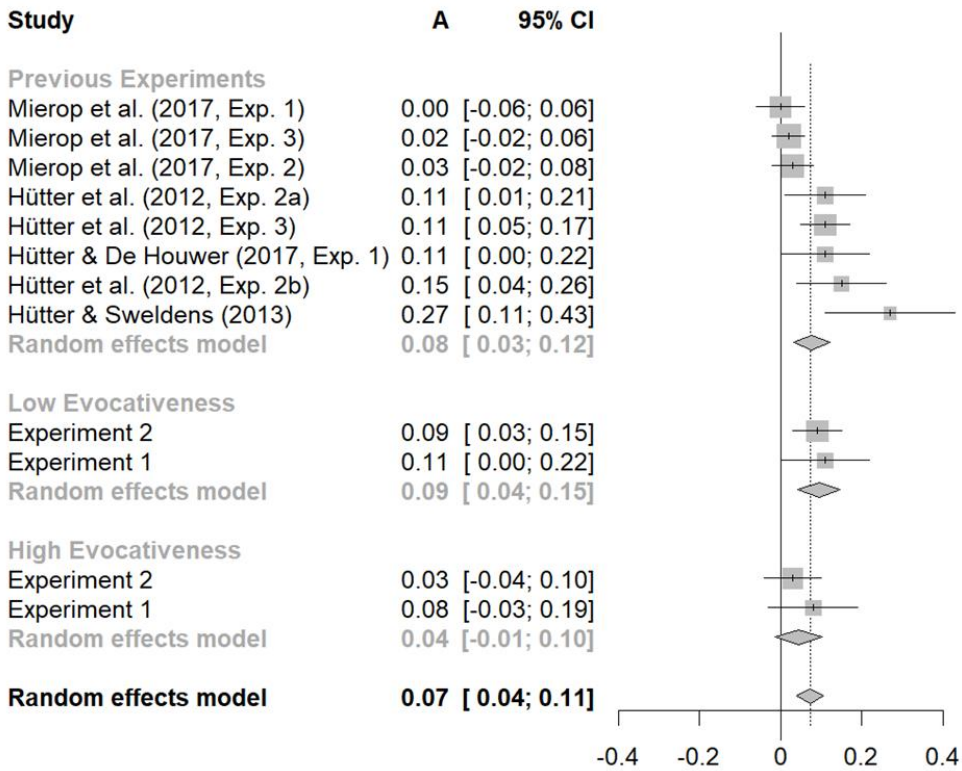


Figure 1. Forest plot depicting the *a*-parameter estimates as well as their respective uncertainties. The estimates are displayed separately for the original/published experiments (top), the low evocativeness conditions (middle), and the high evocativeness (bottom) conditions observed in the two experiments reported here. As can be seen, the low and high evocativeness conditions show values that fall in the middle of the distribution of estimates.

The random effects analysis revealed a mean a-parameter of a = .07, 95% CI [.04, .11]. The experiments also had substantial heterogeneity (I^2 = 56%), indicating that 56% of total variation in the estimates of the a-parameter is due to heterogeneity between experiments rather than chance (Higgins & Thompson, 2002).

Interestingly, the heterogeneity between the experiments was lower when we included the estimates from the two experiments reported here ($I^2 = 56\%$) than when the present experiments were excluded from the meta-analysis ($I^2 = 68\%$). This reduction in heterogeneity implies that the present data fail to provide evidence for assuming that evocativeness has a noticeable effect.

5. General Discussion

To test the hypothesis that memory-independent EC, as measured by the multinomial model, reflects implicit misattribution, we investigated whether it is affected by US evocativeness (Jones et al., 2009). In two high-powered studies, which relied (i) on between and within-participants manipulations of US evocativeness, (ii) on procedures, materials and analyses closely similar to Hütter et al. (2012), and (iii) on various levels of US evocativeness (mild to strong in Exp. 1, and even milder to even stronger in Exp. 2), we consistently found evidence for a memory-independent EC effect (as reflected by the a-parameter) that was not moderated by US evocativeness. In a meta-analysis, we observed that adding the estimates of the a-parameters from the experiments reported here to those from comparable published experiments does not increase the statistical heterogeneity between experiments. We therefore conclude that US evocativeness does not impact the magnitude of the a-parameter, at least under conditions observed here, which are thought to facilitate implicit learning effects.

This lack of support for the IM account is unlikely due to an insufficiently effective manipulation of US evocativeness, as this manipulation (i) was duly pre-tested on evaluative ratings and response times in Experiment 1 and 2, (ii) was further strengthened in Experiment 2, and (iii) impacted the explicit memory parameter m in Experiment 2, thereby replicating the effects on the memory measure reported by Jones and colleagues (2009).

The absence of an effect on the a-parameter may be interpreted in at least four different ways: (i) the a-parameter may not reflect IM-driven EC, (ii) IM may play no role in the EC paradigm considered in the present studies, (iii) IM in the specific paradigm implemented here may be insensitive to evocativeness effects, (iv) evocativeness does not influence the likelihood of IM counter to what has been claimed by Jones et al. (2009). In the remainder of this discussion, we elaborate on why answers to these questions are central to current theorizing in attitude formation and point to important directions for future research.

Whereas experimental studies that decreased participants' ability to encode CS-US pairings in explicit memory found little evidence for memory-independent attitude formation in EC, processing tree studies have supported the view that attitudes may be acquired independently of explicit memory for the CS-US pairings. The fact that an EC effect on the a-parameter was observed only in simultaneous (but not sequential) CS-US settings also seemed consistent with the notion that this parameter captures the operation of an IM process. According to the IM account, people sometimes end up attributing the affective response elicited by the US to the CS. This typically happens when CS and US are presented simultaneously (Jones et al., 2009). Evidence for a significant a-parameter in simultaneous but not in sequential CS-US pairing settings (Hütter & Sweldens, 2013) has been considered strong evidence for the existence of dual-attitude learning (e.g., Sweldens et al., 2014). As discussed in the introduction, however, the manipulation of temporal contiguity does not allow assessing the discriminant validity of the a-parameter (see also Corneille & Stahl, 2018). Consistent with this observation, the present two experiments manipulated US evocativeness, but did not find evidence for the sensitivity of the a-parameter to the latter factor, which is considered another critical moderator of the IM process.

In recent research that crossed the processing tree approach with a cognitive load manipulation at encoding, Mierop and colleagues (2017) expected and found the explicit m-parameter to be largely sensitive to the load manipulation, but failed to obtain evidence for an EC effect on the memory-independent a-parameter. The present research replicates earlier evidence for the existence of EC effects on the memory-independent parameter. Along with Mierop et al. (2017), it also reveals that the m- and a-parameters are differently sensitive to encoding manipulations, which further validates the view that these parameters reflect independent contributions to EC. Yet, it must be noted that the a-parameter is often weak and, perhaps even more critically, that it remains difficult to interpret. This state of affairs is further supported by recent studies showing that EC effects on the a-parameter are observed in instruction-based EC (i.e., in the absence of actual CS-US pairings), and that they vary as a function of manipulations implemented at retrieval (Hütter & De Houwer, 2017). Hence, the a-parameter does not entirely reflect processes occurring at the learning stage. To the extent that it does, we believe it is unlikely to reflect IM based on the present empirical evidence.

Therefore, more research is needed on the conditions under which this parameter is observed and on the processes that may produce it. Importantly, such research should not only focus on the memory-dependency and awareness of learning, but also on other operating conditions and operating principles. Operating conditions refer to features of automaticity, which are manifold and include processing characteristics such as controllability and intentionality (Bargh, 1994; Moors & De Houwer, 2006). These features have played only a minor role in EC research and might even prove more conclusive with regard to the dual-process view (Hütter & Sweldens, in press).

Operating principles refer to the exact nature and the situational constraints of a cognitive process. However, only few theories of evaluative conditioning specify the operating principles of cognitive processes. The IM account by Jones and colleagues (2009, 2010) constitutes a prime example for the concrete conceptualization of a process and its operating principles. Specifically, the IM process is conceptualized to depend on source confusion that is instigated by situational factors such as temporal contiguity of CS and US, US evocativeness, and frequent eye-gaze shifting. Only such clearly specified theories allow for stringent tests of their influence on evaluative learning as performed in the present line of research. Moreover, investigating operating principles might be most informative in identifying the cognitive process(es) contributing to the EC effect in a specific situation as they allow for the strongest tests of discriminant validity. We thus conclude that more such conceptually driven research is needed to advance our understanding of the role of implicit learning processes in the evaluative conditioning paradigm, which is usually considered offering the best support for theorizing on dual-learning of attitudes.

6. Appendix

IAPS Numeric Labels of the US Pictures			
USs -		USs +	
Mild	Intense	Mild	Intense
1080, 1090, 1110,	**1019**, **1030**, 1274,	1450, **1460**, 1603,	**1121**, **1333**, 1340,
2399, 2682, 2691,	**1302**, **1932**, 2095,	1722, **1731**, **1999**,	1500, 1510, **1560**,
2750, **2900(1)**, 3022,	2141, **2220**, **2681**,	**2071**, **2224**, **2304**,	1660, 1900, **1942**,
3220, 5971, 6010,	2683, **2690**, **2692**,	**2345**, **2370**, **2387**,	2299, **2320**, 2344,
6020, **6200**, 6210,	**2694**, 2715, **2722**,	**2388**, 2501, 2655,	2391, **2435**, **2616**,
6211, **6213**, **6230**,	**2780**, **2810**, **3280**,	2791, 2900(2), 4531,	**2620**, **2635**, 2650,
6241, **6242**, **6243**,	3500, 3530, 3550(1),	4536, **4599**, 4601,	**4537**, **4571**, 4653,
6244, **6311**, **6312**,	**5120**, **5130**, **5970**,	**4614**, 4625, **4626**,	**5250**, 5260, 5270
6135, **6370**, 6410,	**5972**, **6000**, **6190**,	**5551**, **5600**, **5621**,	**5390**, 5611, **5622**,
6530, **6570(1)**, **9000**,	6212, 6360, 6561,	**5623**, 5731, **5779**,	5628, **5661**, **5740**,
9001, 9010, 9080,	**6836**, 6838, 9041,	**5830**, 5870, 5875,	**5900**, 7250, 7270,
9110, 9120, **9280**,	9050, **9090**, 9220,	5890, **5891**, 5994,	**7286**, **7352**, 7430,
9340, 9341, **9342**,	**9230**, **9390**, 9452,	7282, 7283, 7325,	7470, **7510**, 7545,
9421, 9440, 9472,	9470, 9471, 9495,	**7350**, 7351, 7360,	7570, **7600**, **7820**,
9530, **9600**, **9620**,	9520, 9611, 9621,	**7580**, 8031, 8032,	8170, **8191**, 8193,
9830, **9910**, **9911**,	9622, 9630, 9635,	**8120**, **8200**, 8280,	8300, **8311**, **8320**,

9913, **9920**	**9700**, **9912**	**8350**, 9531	8497, 8500

Note. IAPS = International Affective Picture System; USs = unconditioned stimuli. USs used in Experiment 2 are in bold font.

Chapter Four

Evaluative influences of CS-US pairings are non-reciprocal

Three experiments examined the reciprocity of evaluative effects following CS-US pairing. In all three experiments, CS evaluations were assimilated to the valence of the US they were paired with (i.e., an evaluative conditioning effect), whereas US evaluations became less extreme (i.e., a US devaluation effect). Of importance, however, US devaluation proved independent of CS-US pairing. Experiment 1 replicated previous evidence for US devaluation: USs were less intensely evaluated after a conditioning procedure as compared to their normative ratings. Experiment 2 controlled for the effect of CS-US pairing: A US devaluation effect of similar magnitude was observed for USs paired with the CSs or presented alone during the conditioning procedure. Experiment 3 indicates that US habituation drives US devaluation: USs presented and evaluated only once were more extremely evaluated than were US paired with CSs or US presented alone during the conditioning procedure, with the latter two US types not differing from each other. Together, these findings suggest that US devaluation is driven by US habituation rather than reciprocal evaluative influences in an associative learning procedure. The theoretical implications of these findings are discussed.

Reference

Mierop, A., Bret, A., Yzerbyt, V., Dumas, R., & Corneille, O. (2018). Evaluative influences of CS-US pairings are non-reciprocal. Manuscript in preparation.

Evaluative influences of CS-US pairings are non-reciprocal

1. Introduction

Evaluations that we hold regarding objects, people and issues along a dimension ranging from negative to positive are a major mediating variable between exposure to information and behavior. In general, people approach what they like and avoid what they dislike, and this is true in many aspects of their life, including social interactions (Greenwald et al., 1998), voting behavior (Bassili, 1993, 1995), or food consumption (Hütter & Sweldens, 2018). Because evaluations are mainly acquired rather than inherited (Rozin, 1982; Rozin & Millman, 1987), it is crucial to understand how they are formed and can be changed.

One of the simplest ways to influence evaluations about one stimulus consists in pairing it with another one. In particular, an initially neutral stimulus generally acquires the valence of a valent stimulus after being paired with it. This effect is referred to as evaluative conditioning (or EC, De Houwer et al., 2001). The initially neutral stimulus is called "conditioned stimulus" (or CS) because its acquisition of valence is conditional on its pairing with the valent stimulus. The valent stimulus is called "unconditioned stimulus" (or US) because it elicits an evaluative response prior to its pairing with the CS.

In sum, CSs are conditioned following their pairing with USs. But are US evaluations also influenced by their pairings with CSs? The current research examined this question. We first provide a short state-of-the-art discussion of this question and then report three experimental studies that addressed it.

1.1 Is there evidence that the CS influences the US evaluation and why does that question matter?

That the US elicits an evaluation prior to its pairing with the CS does not imply that its evaluation is not to be altered as a result of that pairing. The US, too, may be conditioned (i.e., it may also be a conditioned stimulus). Consistent with this view, (De Houwer, Baeyens, Vansteenwegen, & Eelen, 2000) observed that USs elicit milder evaluations after been paired with CSs. In their experiment, these authors had their participants rate 70 face stimuli. Then, they selected, for each participant, the eight most neutral and the four most positive and four most negative stimuli to design the CS-US pairs. They observed the typical evaluative conditioning effect on the CSs: CS ratings were assimilated to the valence of the US they were paired with. More critical to the present research, they also found a "US devaluation" effect: the USs were evaluated less extremely after being paired with the CSs. Hence, it seems that the two components in the CS-US pair influenced each other in opposite directions, resulting in more polarized evaluations of the CSs (i.e., an evaluative conditioning effect) and milder evaluations of the USs (what we will refer to here as a 'US devaluation" effect).

The US devaluation effect has important practical and theoretical implications. One practical implication is clinical in nature: if people acquire a strong dislike for certain category of stimuli (e.g., spiders), is there any advantage in actively pairing these with neutral stimuli, rather than presenting them in isolation?

At the theoretical level, the US devaluation question is relevant to associative versus propositional attitude learning views. As compared to propositions, associations are "unqualified" in nature, and as such do not contain structural content about the CS-US relation (De Houwer, 2009; Shanks, 2007). Hence, associations may not be expected to produce non-reciprocal evaluative influences. Rather, they may produce a reciprocal influence since the CS and the US are bound and should either co-activate each other through excitatory links or interfere with each other through inhibitory links. Of course, this analysis does not imply that US-to-CS and CS-to-US influences should be of same magnitude. Yet, finding an exclusively US-to-CS influence may invite to qualify associative learning views (see also the general discussion).

A second question of theoretical interest is why US devaluation is found when EC effects are typically resistant to extinction (e.g., Baeyens et al., 1988; Dwyer, Jarratt, & Dick, 2007)? That is, why pairing a valent stimulus with a neutral stimulus makes its evaluation milder (i.e., US devaluation), whereas CS evaluations acquired in an EC procedure remain typically stable when the now-valent CS is presented without the US after the conditioning phase (i.e., resistance to extinction)?

In order to address these questions, it is important to proceed first to a closer examination of the processes underlying the US devaluation effects. In particular, two processes may account for US devaluation that are *independent of any effect of CS-US pairing*. First, US devaluation may reflect a regression to the mean effect. If a variable is extreme upon its first measurement, it will tend to regress to the average on a subsequent measurement (Healy & Goldstein, 1978). When USs are selected for each participant based on their most extreme evaluations prior to CS-US pairing, not only the 'true' USs value but also their measurement errors are likely to be extreme. As a result of the redistribution of random measurement errors on the second (i.e., post CS-US pairing) measurement, USs evaluations are expected to become less extreme, even though their true value remains unchanged. Second, US devaluation may be driven by affective habituation. Because CS-US pairing is confounded with exposure to the US, US devaluation may reflect "reductions in stimulus-evoked affective reactions as a result of previous exposure" (Leventhal, Martin, Seals, Tapia, & Rehm, 2007). Neither of these processes would have anything to do with CS-US pairing *per se*.

The present experiments aimed at testing (1) the robustness of the US devaluation effect reported by De Houwer et al. (2000) in procedures devoid of regression to the mean effect, (2) whether there is a CS-US pairing contribution to this effect and, if there is not, (3) for the contribution of the mere presentation of the US.

1.2 Overview of the Experiments

Experiment 1 conceptually replicated De Houwer et al. (2000) by measuring evaluative changes in both CSs and USs after CS-US pairing. Experiment 2 examined the contribution of CS-US pairings in US devaluation by introducing non-associative control USs that were presented but not paired during the conditioning phase. Experiment 3 further addressed the role of habituation by comparing the evaluation of USs that were never presented, USs that were paired with CSs, and US that were presented alone in the learning procedure. All three experiments relied on normative ratings for the selection of the USs. Because normative ratings come from many individual evaluations of the same stimuli set, measurement errors are randomly distributed such that regression to the mean effects cannot be conducive to a systematic devaluation effect.

1.3 Statistical power and sample size

De Houwer and colleagues (2000) reported a – very large – US devaluation effect of Cohen’s d = 2.5. We decided to draw on the average EC effect of Cohen’s d = .52 (Hofmann et al., 2010) to determine our samples size. If the US devaluation is as strong as the EC effect (if CSs and USs change to a same degree), we need 31 pairs of observations to achieve a high statistical power $(1 - \beta) = .8$ with a two-tailed Type I error $\alpha = .05$. With a minimum of 89 pairs of observation (Experiment 3), we secured enough sensitivity for observing a devaluation effect as small as d = .3 with a high statistical power $(1 - \beta) = .8$.

2. Experiment 1

Experiment 1 allowed testing the presence of the US devaluation effect reported by De Houwer et al. (2000) in a procedure devoid of any regression to the mean effect. This experiment was originally aimed at exploring unrelated hypotheses. Therefore, we collected additional measures (i.e., inter-individual measures) that are not reported here. The data from the experiment as well as the R script created to manage, analyze, and represent visually the data are available on Open Science Framework (OSF link: osf.io/82mcs).

2.1 Method

2.1.1 *Participants*

136 participants (M_{age} = 24.32, SD_{age} = 5.99, 99 women) from a large European university participated without compensation. They were contacted by email and invited to take part in an online study. The experiment was programmed using the 'Testable.org' platform.

2.1.2 *Procedure and materials*

The procedure involved three phases. In the first - conditioning - phase, eight CS-US pairs were displayed on the screen. Four CSs were presented with positive USs and four CSs were presented with negative USs. The specific CS-US valence pairings were created randomly for each participant. Each pair was presented eight times, in a random order, for 1000ms. The Inter-Trial-Interval (ITI) was of 500ms. CSs and USs were borrowed from Pleyers and colleagues (2007). CSs depicted consumer goods and USs were affective pictures from the International Affective Picture System (IAPS, Lang et al., 1997) representing people in happy or sad situations (see Appendix). In the second – CS evaluation - phase, participants reported their global feelings towards each CS on a 9 point-scale (1 = very negative; 9 = very positive). In the third – US evaluation - phase, participants rated each US on a scale from 1 (very unpleasant) to 9 (very pleasant).

2.2 Results

2.2.1 *Analytical strategy*

For each participant, we computed both for the CSs and for the USs the average difference in ratings between post-conditioning ratings and normative ratings. Normative ratings for the CSs were collected in a pre-test (see Pleyers et al., 2007) and normative ratings for the USs consisted in IAPS ratings. The mean differences in evaluations were averaged, separately for each US valence (positive and negative). The data were analyzed using the 'aov' and 'anovaBF' functions in R (R Development Core Team, 2017; from the R package 'BayesFactor', Morey, Rouder, & Jamil, 2015) for the frequentist and Bayesian analyses respectively. We report the Bayes factors associated with the model comparison made in the frequentist analyses. An augmented model containing the tested factor was compared to a constrained model that did not contain the tested factor. The Bayes factors in favor of the alternative hypothesis (or BF10) are presented when the conventional p-value of .05 is encountered. The Bayes factors in favor of the null hypothesis (or BF01) are reported when the p-value is above this threshold.

2.2.2 *CSs*

Changes in CS evaluation were submitted to a repeated-measures ANOVA with US valence (positive vs. negative) as a within-participant factor. The main effect of US valence was significant, $F(1, 135) = 104.3$, $p < .001$, $\eta^2_p = .43$, BF10 > 1000: CSs paired with positive CSs (i.e., CSs+) became more positive ($M = 1.09$, $SD = 1.43$, $t(135) = 8.88$, $p < .001$) whereas CSs paired with negative USs (i.e., CSs-) became more negative ($M = -.76$, $SD = 1.50$, $t(135) = -5.88$, $p < .001$). Hence, the data revealed the typical evaluative conditioning effect (see Figure 1).

2.2.3 ***USs***

Changes in US evaluation were submitted to a repeated-measures ANOVA with US valence (positive vs. negative) as a within-participant factor. The main effect of US valence was significant, indicating that USs evaluations became more neutral after the pairing (i.e., a US devaluation effect), $F(1, 135) = 22.15$, $p < .001$, $\eta^2_p = .13$, BF10 > 1000. Looking at each valence separately, positive USs became more negative ($M = -.61$, $SD = 1.26$, $t(135) = -5.70$, $p < .001$), whereas negative USs tended to become more positive ($M = .16$, $SD = 1.15$, $t(135) = 1.59$, $p = .11$; see Figure 1).

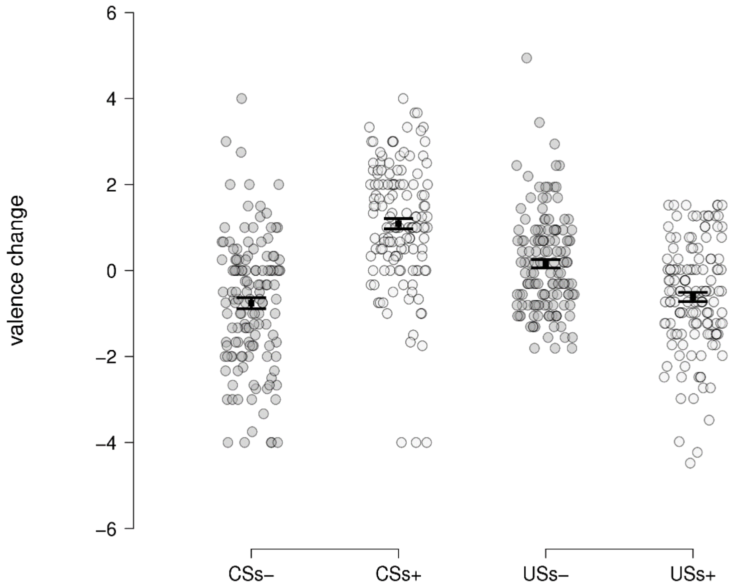


Figure 1. Change in evaluation for positive and negative CSs and USs. Dots represent individual responses. Filled squares represent observed means. Error bars represent standard errors of means. CSs- = negatively conditioned stimuli. CSs+ = positively conditioned stimuli. USs- = negative unconditioned stimuli. USs+ = positive unconditioned stimuli.

2.3 Discussion

In Experiment 1, CSs were more valent after than before the EC procedure (i.e., an evaluative conditioning effect emerged). Conversely, USs elicited less extreme evaluations after than before CS-US pairing (i.e., the US devaluation effect was replicated). Because we used normative ratings as a baseline, it is unlikely that a regression to the mean effect drove US devaluation. Instead, US devaluation may be due to the CS-US pairing or to the mere presentation of the USs. Of note, US devaluation was significant only for the positive USs. One should therefore remain cautious about the interpretation of this result as it may be due to the reliance on different samples for normative and experimental ratings. Collecting evaluations before and after CS-US pairings within the same sample of participants allows overcoming this interpretational issue. This is the procedure used in Experiment 2.

3. Experiment 2

The role of the CS-US paring was examined in Experiment 2 by including control US-alone stimuli in the conditioning procedure. As a further asset, Experiment 2 relied on idiosyncratic rather than normative US evaluations. In this case, each participant served as her own control, although USs were still selected on the basis of prior normative ratings (therefore precluding regression to the mean effects). We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the experiment. The raw data from the experiment as well as the R script created to manage, analyze, and represent visually the data are available on Open Science Framework (OSF link: osf.io/82mcs).

3.1 Method

3.1.1 *Participants*

We collected data on 163 participants (M_{age} = 33.74, SD_{age} = 10.39, 78 women) on the online platform ‘Prolific academic’. Participants received a compensation of 0.6 £ for their participation.

3.1.2 *Procedure and materials*

Participants first rated eight CSs (Pleyers et al., 2007) and 16 USs (see Appendix) using a scale from 1 (very negative) to 9 (very positive). Eight randomly selected USs always appeared alone (US-alone trials) and the remaining eight USs were systematically paired with CSs (US-paired trials) during the conditioning procedure. For the US-alone trials, the US appeared alone for 1000ms. For the US-paired trials, the US was simultaneously paired with a CS for 1000ms. All trials were randomly ordered and separated by an ITI of 500ms. After the conditioning procedure, participants rated the CSs and USs anew.

3.2 Results

3.2.1 *Analytical strategy*

We computed the difference between post- and pre-ratings for the CSs and USs at the stimulus level for each participant. A positive value reflects more positive evaluations at post-test. We then averaged evaluative change scores by participants as a function of US valence and pairing type. As in Experiment 1, we used frequentist and Bayesian repeated-measures ANOVA to analyze the data.

3.2.2 *CSs*

Changes in CS evaluation were submitted to a repeated-measures ANOVA with US valence (positive vs. negative) as a within-subject factor. This analysis revealed a main effect of US valence, $F(1, 162) = 25.37$, $p = < .001$, $\eta^2_p = .13$, BF10 > 1000. As expected, CSs+ became more positive ($M = .39$, $SD = .98$, $t(162) = 5.03$, $p < .001$), whereas CSs- became more negative ($M = -.27$, $SD = 1.17$, $t(162) = -2.96$, $p < .001$). Hence, an EC effect was found (see Figure 2).

3.2.3 *USs*

Changes in US evaluative ratings were submitted to a repeated-measures ANOVA with US valence (positive vs. negative) and pairing type (alone vs. paired) as within-subject factors. We observed a main effect of US valence, $F(1, 162) = 5.61$, $p = .019$, $\eta^2_p = .03$, BF10 = 8.20. As for Experiment 1, positive USs (M = -.24, SD =.87, $t(162) = -3.99$, $p < .001$) changed in valence whereas negative USs did not (M = -.02, SD = 1.12, $t(162) = -.318$, $p = .75$). We did not observe a US valence × pairing type interaction, $F(1, 162) = .55$, $p = .459$, $\eta^2_p = .00$, BF01 = 5.81. As depicted in Figure 2, the US devaluation effect did not differ between the US-alone trials and the US-paired trials. No main effect of pairing type was observed either (F < 1).

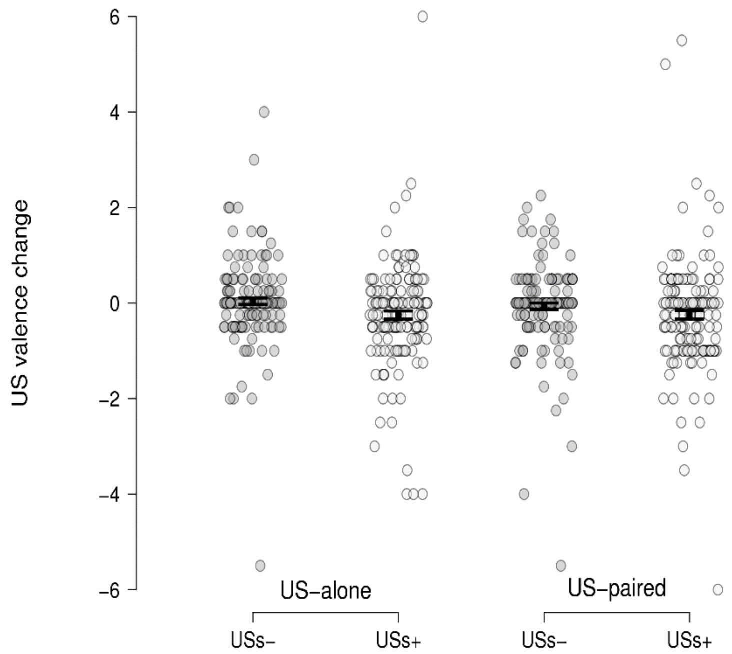


Figure 2. Change in evaluations of the USs as a function of US valence and pairing type. Dots represent individual responses. Filled squares represent observed means. Error bars represent standard errors of means. USs- = negative unconditioned stimuli. USs+ = positive unconditioned stimuli.

3.3 Discussion

Experiment 2 again found evidence for both an EC effect and a US devaluation effect. It provides no evidence, however, that CS-US pairing effect contributed to US devaluation: US presented alone and US paired with CSs decreased in valence to a similar extent. The Bayesian analysis suggests that the null effect of pairing is over five times more likely than an effect of pairing in the devaluation effect. Once again, the US devaluation was observed only for the positive USs. This time, this cannot be due to a difference in the samples investigated (i.e., normative versus experimental). Instead, this imbalance might suggest an asymmetric affective habituation effect, with larger habituation to positively than negatively charged stimuli.

4. Experiment 3

Experiment 3 was designed to investigate further the role of CS-US pairing in US devaluation. In addition, this experiment provides a formal test for the contribution of US habituation. This was achieved by comparing evaluations for USs that were presented and rated only once *versus* USs that were either paired with CSs or presented alone in the conditioning procedure. If US habituation plays a role, ratings should be less extreme for US-alone and US-paired stimuli than for US presented and rated only once. We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the experiment. The raw data from the experiment as well as the R script created to manage, analyze and represent visually the data are available on Open Science Framework (OSF link: osf.io/82mcs).

4.1 Method

4.1.1 *Participants*

We collected data on 82 participants (M_{age} = 22.90, SD_{age} = 5.11, 56 women) on the online platform 'Prolific academic'. Participants received a compensation of 0.6 £ for their participation.

4.1.2 *Procedure and materials*

A total of 24 USs were used (see Appendix), 12 positive and 12 negative. During the conditioning procedure, eight were never presented; eight were presented alone; eight were paired with CSs. The US-alone and the US-paired trials were presented in random order for 1000ms each, with an ITI of 500ms. Participants then rated the CSs and the USs using an evaluative scale from 1 (very negative) to 9 (very positive).

4.2 Results

4.2.1 *Analytical strategy*

We averaged CS and US evaluative ratings within each US valence and pairing type for each participant. As for the previous experiments, we used frequentist and Bayesian repeated-measures ANOVA to analyze the data.

4.2.2 *CSs*

Evaluations of the CSs were submitted to a repeated-measures ANOVA with US valence (positive vs. negative) as a within-subjects factor. This analysis revealed a main effect of US valence, $F(1, 81) = 38.06$, $p < .001$, $\eta^2_p = .31$, BF10 > 1000: CSs+ were evaluated more positively ($M = 6.01$, $SD = 1.29$) than CSs- ($M = 4.82$, $SD = 1.38$), therefore confirming an EC effect (see Figure 3)

4.2.3 *USs*

Evaluations of the USs were submitted to a repeated-measures ANOVA with US valence (positive vs. negative) and pairing type (US-never vs. US-alone vs. US-paired) as within-participant factors. This analysis revealed a main effect of US valence, $F(1, 81) = 1032$, $p < .001$, $\eta^2_p = .93$, BF10 > 1000: positive USs ($M = 7.31$, $SD = 1.16$) were evaluated more positively than negative USs ($M = 1.99$, $SD = .91$). This analysis further revealed a US valence × pairing type interaction, $F(2, 162) = 10.77$, $p < .001$, $\eta^2_p = .11$, BF10 > 1000 (see Figure 3). Specifically, the difference in US ratings was larger for US-never trials (M = 5.70, SD = 1.78) as compared to both US-alone trials (M = 5.09, SD = 1.75; $F(1, 81) = 15.32$, $p < .001$, $\eta^2_p = .15$, BF10 > 1000) and US-paired trials (M = 5.14, SD = 1.51; $F(1, 81) = 15.07$, $p < .001$, $\eta^2_p = .15$, BF10 > 1000. There was no difference between the latter two conditions, $F(1, 81) = .125$, $p = .724$, $\eta^2_p = .00$, BF01 = 5.36. Looking at habituation effects within each US valence, evidence was now found both for the positive US and negative USs. Specifically, ratings between learning conditions differed both for positive USs, $F(2, 162) = 3.51$, $p = .03$, $\eta^2_p = .03$, BF10 = 1.92, and negative USs, $F(2, 162) = 12.39$, $p < .001$, $\eta^2_p = .12$, BF10 > 1000.

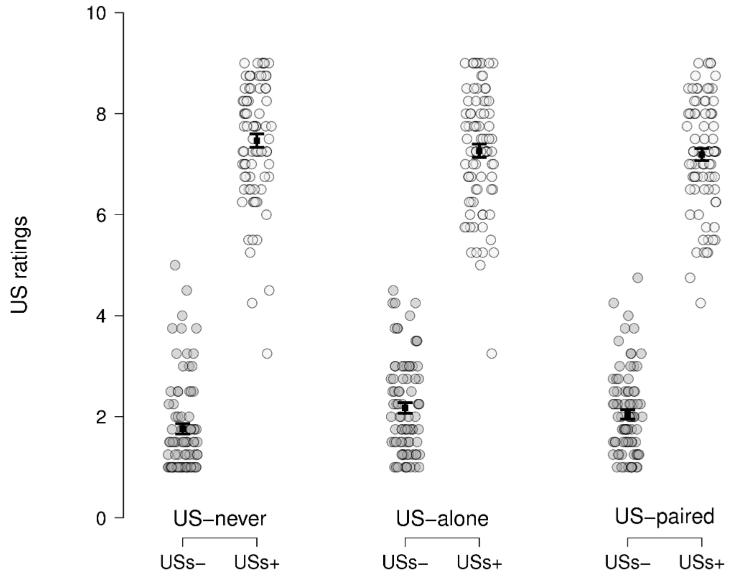


Figure 3. US ratings as a function of US valence and pairing type. Dots represent individual responses. Filled squares represent observed means. Error bars represent standard errors of means. USs- = negative unconditioned stimuli. USs+ = positive unconditioned stimuli.

4.3 Discussion

In Experiment 3, USs presented alone or with a CS during the conditioning procedure were evaluated less intensely (and equally so) than were USs never presented before they were evaluated at the end of the experiment. Hence, the present study again finds no evidence for the contribution of CS-US pairing in the US devaluation effect. In addition, it provides evidence for the contribution of the mere US presentation to this effect. Of note, the US devaluation effect was found this time for both the positive and negative USs.

5. General discussion

In this research, we examined the reciprocity of evaluative changes following CS-US pairings. An evaluative conditioning effect was systematically observed, which confirms the influence of the US in the evaluation of the CS after pairing these together. Consistent with De Houwer et al. (2000), evidence for a US devaluation effect was systematically found, with reduced evaluations of the USs following their pairings with the CS. Of importance, however, various control conditions suggest that this devaluation effect is immune to any CS-US pairing influence. Specifically, in Experiment 2 and 3, the US devaluation effect was observed to the same extent for USs paired with CSs and USs presented alone during the conditioning procedure. Hence, CS pairing did not contribute to US devaluation. Rather, the devaluation effect is likely driven by an affective habituation effect: USs presented (alone or with a CS) during the conditioning procedure were evaluated less intensely (and to the same extent so) than were USs never presented before they were rated at the end of the study. Across studies, habituation was systematically observed for the positive USs. As to the negative USs, there was a trend for habituation in Experiment 1 and the effect was clearly significant in Experiment 3, but it was absent in Experiment 2. Although this suggests that people may be more likely to habituate to positive than to negative USs, more research is needed on this front.

These findings suggest that evaluative changes elicited by CS-US pairings are nonreciprocal, in the sense that pairing the US with a CS has no demonstrated influence on the US evaluation over and above that elicited by US habituation. The one-sided evaluative effect of the CS-US association seems problematic for dual-models of attitudes that rely on associative assumptions. Specifically, it is unclear why unqualified associations with stimuli should show nonreciprocal effects (see also below). One possibility is that CSs, being neutral, elicit no response, the latter of which would be needed for establishing evaluative effects. Whether an absence of valent response should be equated to the presence of a neutral response is an interesting question. To the extent that non-valent stimuli are evidenced by neutral evaluative ratings, the answer to this question is affirmative: neutral stimuli, clearly, do result in evaluative responses. It is therefore important to specify the nature of observable responses elicited by neutral stimuli that would or would not be relevant to associative or dual-learning models of attitude. The latter specification would help to constrain further these models.

Alternatively, CS-to-US effects may have been too small to be detected in Experiment 2 and 3, or may apply to other pairing or evaluative measurement paradigms that remain to be uncovered. Providing evidence for the inexistence of CS-to-US effects in EC procedures is formally impossible. If we cautiously assume, however, that effects of CS-US pairings are non-reciprocal, then this would suggest that the evaluative effect of CS-US linkage is sensitive to a reciprocal qualifier, one that goes from the US to the CS only, at least by default. In turn, this qualified (in this case, reciprocal) CS-US linkage is compatible with propositional accounts of attitude learning (e.g., Fiedler & Unkelbach, 2011; Mitchell et al., 2009; Zanon, De Houwer, & Gast, 2012). As just discussed, it would seem less consistent with an associative learning account, commonly defined as involving a low-level process that automatically registers co-occurrences between stimuli independent of their validity and relational meaning (for a recent discussion, see Corneille & Stahl, 2018). We believe this issue should encourage to constrain further dual-models of attitude learning with the extra assumptions (1) that weakly valent and weakly arousing stimuli elicit no evaluative response relevant to these models, and that (2) evaluative effects are contingent on the active production of evaluative responses. Turning to propositional models, they may be further tested in experiments that reverse the assumed US-to-CS relational qualifier, for instance by instructing participants to pay more attention to the meaning of the CS-to-US relation.

Appendix

IAPS Numeric Labels of the US Pictures

USs -	USs +
6312, 2750, 6360, 6561,	4608, 4700, 8200, 8460,
2141, 2900, 6315, 6510,	2550, 4603, 4641, 8120,
6210, 6312, 6550, 6571	1603, 2501, 8162, 7325

Note. IAPS = International Affective Picture System; USs = unconditioned stimuli. All USs were used in Experiment 3. USs from the first two rows were used in Experiment 2. USs from the first row were used in Experiment 1.

Chapter Five

Does Extinction in Evaluative Conditioning Require Producing Evaluative Responses During Extinction Training?

Two high-powered experiments examined the role of evaluative response production in the extinction of evaluative conditioning (EC) by positioning EC in the procedural and conceptual framework of classical conditioning (CC). According to Rescorla's (1993, 1997) response theory of extinction, the occurrence of a response is required for that response to extinguish in CC. And the more frequent this response, the more effective the extinction. Experiment 1 used three extinction conditions following response acquisition in an EC procedure: evaluative responses were measured only after extinction; after acquisition and after extinction; or were continuously measured after acquisition, during extinction and after extinction. Consistent with the response theory of extinction, we predicted that extinction of EC would be the highest in the third condition. Experiment 2 was aimed at further facilitating extinction of EC by encouraging participants to experience that their evaluation may change over the course of the experiment. To this end, half of the participants completed pre- and post-acquisition ratings prior to practicing continuous response expression in the extinction phase. Contrary to our predictions, no extinction of EC was observed in either of these experiments. We conclude that Rescorla’s (1993, 1997) response theory of extinction may not apply to EC and discuss the theoretical implications of this finding.

Reference

Mierop, A., Molet, M., & Corneille, O. (2018). Does Extinction in Evaluative Conditioning Require Producing Evaluative Responses During Extinction Training? Manuscript in revision, *Cognition and Emotion*.

1. Introduction

Evaluative conditioning (EC) is defined as a change in the evaluation of a conditioned stimulus (CS) following its pairing with an affective unconditioned stimulus (US) (e.g., an increased evaluation for a neutral image after it has been paired with a pleasant image). As evidence accumulated, it has become increasingly clear that EC shares some, but not all, functional characteristics of classical conditioning (CC). For instance, similar to CC, counter-conditioning (Kerkhof, Vansteenwegen, Baeyens, & Hermans, 2011), latent inhibition (Díaz & De la Casa, 2002), sensory preconditioning (Hammerl & Grabitz, 1996), and US revaluation (Walther et al., 2009) have been observed in EC research. By contrast, for example, blocking (Beckers, De Vicq, & Baeyens, 2009), occasion setting (Frank Baeyens, Crombez, De Houwer, & Eelen, 1996), and overshadowing (Dwyer et al., 2007) have not been detected for EC.

In the EC literature, it has been a recurrent question whether EC shares a well-established functional characteristic of CC: extinction. That is, whether presenting repeatedly a CS alone after it has acquired a valence through its pairing with a US attenuates its evaluation. It seems that, contrary to CC (Delamater, 2004), EC does not extinguish (Baeyens, Díaz, & Ruiz, 2005; Frank Baeyens et al., 1988; Blechert, Michael, Williams, Purkis, & Wilhelm, 2008; De Houwer, Baeyens, Vansteenwegen, & Eelen, 2000b; Dwyer et al., 2007; Engelhard, Leer, Lange, & Olatunji, 2014; Hermans et al., 2005; Kerkhof, Vansteenwegen, Baeyens, & Hermans, 2009; Vansteenwegen, Francken, Vervliet, De Clercq, & Eelen, 2006). A closer look at the literature however suggests that there is no clear-cut evidence that EC does not extinguish.

In a recent meta-analysis, however, Hofmann and colleagues (2010) found that EC extinction was observed, but exclusively in the subset of studies making use of CS evaluative ratings both after the conditioning and the extinction procedures. This meta-analysis highlights the need of sufficient statistic sensitivity since the individual studies even within this subset failed to obtain significant EC extinction. Furthermore, it suggests that CS ratings between the EC and extinction procedures might be a necessary condition for EC extinction. This is consistent with the findings of Gawronski and colleagues (2015) who found partial extinction when they used a within-subject design (i.e., comparing CSs ratings after acquisition to CSs ratings after extinction in the same group) (Experiments 1 and 3) but not when they used a between-subject design (i.e., comparing the CSs ratings from one group that evaluated the CSs after acquisition to the CSs ratings from another group that evaluated the CSs after extinction) (Experiment 2). They furthermore observed partial extinction on a direct evaluative measure but not on an indirect one.

On a related though distinct note, in two experiments, Lipp, Oughton, and LeLievre (2003) relied on an online measure of CS evaluations across the extinction procedure (i.e., once every four CS presentation). They observed that the CSs evaluations declined progressively across the extinction procedure towards the neutral level. The authors suggested that the valence of the CSs has to be rated repeatedly during the extinction procedure to observe extinction. Two additional experiments replicated EC extinction when using multiple CS ratings during the extinction procedure (Lipp & Purkis, 2006). Interestingly, in these studies, a different evaluative mode for the final test did not show a persistence of the extinction (Lipp et al., 2003; Lipp & Purkis, 2006).

In sum, extinction of EC effects would seem to depend on (1) the presence of explicit CS evaluations between the acquisition and extinction procedures or (2) the presence of multiple evaluations of CSs in the extinction procedure, provided these evaluations are produced in a unique response mode.

To date, this conditional EC extinction has been interpreted as reflecting judgmental instead of learning effects (Aust et al., in press; Gawronski, Gast, & De Houwer, 2015; Lipp & Purkis, 2006). More specifically, drawing on research on causal judgment (i.e., Collins & Shanks, 2002; Matute, Vegas, & De Marez, 2002), Lipp and Purkis (2006) hypothesized that CS evaluations may reflect judgments based either on the most recent information encountered about the CS or based on information integrating the whole learning history about the CS. When CSs are evaluated only once after the whole learning history, their evaluations reflect the integration of all episodes about the CSs. When CSs are evaluated multiple times throughout the experiment, however (e.g., before and after the extinction procedure), their evaluations may reflect only the most recently acquired information (e.g., that acquired during the extinction procedure only).

Whereas current accounts of extinction of EC point to strategies used at the judgment stage, another account (i.e., Rescorla’s response theory of extinction) points to a change in the CS-US representation. Below, we explain the tenets of this theory, commonly applied to extinction in classical conditioning. Next, we report two experiments aimed at testing it in the context of EC. In doing so, the present research advances our understanding of the mechanisms underlying the extinction of EC. It also delivers further insights into the functional correspondence between EC and classical conditioning.

1.1 The Response Theory of Extinction

The dominant idea in the classical conditioning (i.e., CC) literature, emerging from different theories, is that extinction does not reflect an “unlearning” of acquired CS-US mental links but rather reflects a “new learning” of an inhibitory CS-US association that interferes with the activation of the conditioned response (CR) (e.g., Bouton, 1993; Hull, 1943; Pavlov, 1927).

Of particular interest here is Rescorla's (1993; 1997) response theory of extinction. According to this theory, two types of association between a stimulus and a response are involved: one is excitatory and the other inhibitory. More specifically, an excitatory link is formed between the CS and a response elicited by a US during the course of conditioning.

Therefore, conditioned responding depends on the ability of the CS to activate a representation of the response. During extinction training, the repeated presence of the CS in the absence of the US progressively boosts the inhibitory link between the CS and the response, which means that inhibition becomes larger than excitation. As soon as this is the case, an extinction of the CR is observed. This view has received support in the CC literature (e.g., Colwill, 1991; Delamater, 1996; Rescorla, 2001). Related results have also been reported for instrumental conditioning (e.g., Bouton, Todd, Vurbic, & Winterbauer, 2011; Bouton, Trask, & Carranza-Jasso, 2016).

Importantly, the response theory of extinction posits that extinction does not require the presentation of a non-reinforced CS *per se*, but rather the production of a conditioned response that is not reinforced by a US. We explain the reasons for this point, as well as its implications below.

1.2 Relevance of a theory of CC to EC

As mentioned above, the current question should advance our understanding of extinction in EC, and also inform us on the functional correspondence between EC and CC effects. EC is best conceptualized as an effect and should not be confounded with CC. However, this does not imply that we should stop looking at EC in the framework of CC (e.g., Olson & Fazio, 2001). As a matter of fact, looking at whether EC is a form of CC has generated an important number of theoretical questions that helped us to better understand EC effects (e.g., De Houwer, 2011; Field, 2000; Hammerl & Grabitz, 2000; Kim et al., 2015). Within this perspective, we focus on two important questions about the extinction of EC effects. First, are extinction procedures similar in CC and EC? Second, is the supposed general mechanism causing extinction in CC at play in EC?

In CC, the CR is automatically and continuously produced by the autonomous nervous system during extinction training: the CS comes to elicit a reflex-like learned response that eventually extinguishes during extinction training (Pavlov, 1927; Rescorla, 1988). In contrast, however, it is less clear whether a CR is triggered during extinction in EC paradigms. Admittedly, attitude research suggests that valent stimuli (including CSs) easily give rise to evaluative responses. It is thought, for instance, that the affective priming effect reflects automatic evaluative responses (Fazio, Sanbonmatsu, Powell, & Kardes, 1986). Research has also shown, however, that mild evaluative stimuli elicit relatively milder physiological responses (e.g., Bernat, Patrick, Benning, & Tellegen, 2006). And, more importantly, the production of evaluative responses may rely on higher-level processes, such as evaluative goals (e.g., Everaert, Spruyt, & De Houwer, 2013; Spruyt, Houwer, & Hermans, 2009).

Consistent with the response theory of extinction, we reasoned that a possible lack of evaluative responding during extinction training in EC studies could have been responsible for the lack of extinction. If so, and assuming acquired evaluative responses in EC are sensitive to the same extinction mechanisms as those underlying the extinction of automatic responses in CC, ensuring an evaluation of the CSs during extinction training (through continuous CS ratings) may be conducive to extinction. This theory would also predict that the frequency of response production plays an important role in EC extinction: the more frequent the non-reinforced response production and the more effective the extinction procedure. In the present research, we manipulated for the first time the frequency of response production, and we also for the first time created a continuous response production setting, the latter of which, based on the response production theory, may maximize EC extinction.

1.3 How consistent is the Response Production Theory with existant evidence on EC extinction?

Assuming that an evaluative conditioned response is not automatically produced upon the mere perception of a CS, the response theory of extinction is consistent with the lack of EC extinction when disposing of no other CS evaluation than a final one after the extinction procedure (Gawronski et al., 2015; Hofmann et al., 2010). Indeed, in this case no response is produced at extinction.

The Response Theory of Extinction also explains the (weak) presence of extinction in studies that involved a CS evaluation directly after the learning phase. In this case, a response is produced for at least one of the many extinction trials (i.e., the first one).

This theory is also consistent with the fact that extinction is observed only with direct evaluative measures (Gawronski et al., 2015). Because, by definition, indirect evaluative measures do not require the direct production of an evaluation of the CS, they do not allow for response inhibition.

It is also consistent with progressive EC extinction across the extinction procedure when evaluative conditioned responses are repeatedly produced. This is because each non-reinforced production of conditioned response should strengthen the inhibitory link between the CS and the conditioned response, which should therefore become progressively less intense.

Finally, the Response Theory of Extinction has difficulties in explaining that EC extinction does not generalize to another response mode (Lipp et al., 2003). Because this theory points to representational effects, these effects should be expected to generalize across measurement modes. Of note, however, Lipp et al. (2003) explained away this particular finding by pointing to a possible renewal of the CS-US association provided a change in context.

1.4 The present research

Two experiments manipulated for the first time the frequency of the evaluative conditioned response production at extinction as a randomly assigned factor (see Figure 1). In a first experiment, we compared the two traditional extinction procedures (with and without post-acquisition ratings) to an extinction procedure adding continuous evaluative ratings. The latter condition ensured a continuous production of conditioned responses throughout the extinction phase. At a functional level of analysis, we were interested in whether the mere presence of CS evaluations before the extinction procedure is sufficient to produce EC extinction. We were furthermore interested in examining the added value of the continuous production of evaluative conditioned responses throughout the whole extinction procedure.

In a second experiment, we examined whether asking participants to rate the to-be-conditioned stimuli prior to acquisition would facilitate extinction. We reasoned that extinction could be facilitated when participants are encouraged to experience that their evaluations may change over the course of the experiment, which is clearly the case when reporting an evaluation before versus directly after acquisition. Combining response expression with such primer may represent an ideal case for detecting extinction.

2. Experiment 1

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the experiment. The raw data from the experiment as well as the R script created to manage, analyze and visually represent the data are publicly available on Open Science Framework (OSF link: osf.io/ky9tn).

2.1 Method

2.1.1 *Participants*

The sample size of this experiment was estimated on the basis of the effect size of the extinction effect estimated in the meta-analysis of Hofmann and colleagues (2010). With an extinction (i.e., within-participants differential effect between post-extinction and post-acquisition CS ratings) effect size of Cohen's d = .322, 106 participants are required to obtain a statistical power of (1 - β) = .95, with an alpha of .05 (one-tailed given the unilateral prediction: extinction reduces the conditioned evaluative responses).

We collected data on 260 participants (M_{age} = 33.74, SD_{age} = 10.39, 102 women) on the online platform 'Prolific academic'. Participants were compensated of 0.6 £ for their participation. Among these participants, 179 were assigned to one of the two experimental conditions containing evaluative ratings both after the acquisition phase and the extinction phase and therefore met the within-participants extinction design. The latter sample size allows disposing of 99.5% chance to replicate a significant extinction effect (unilateral alpha inferior or equal to .05) assuming a 'true' effect size of Cohen's d = .322 (as reported by Hofmann and colleagues, 2010).

With regard to the potential moderating effect of the response frequency, the sample size of 260 participants allowed to detect a between-participants effect as small as of Cohen's $f^2 = .06$, which represents a small-to-medium effect size according to Cohen's (1988) norms, with a statistical power of $(1 - \beta) = .95$ and a two-tailed alpha of .05.

2.1.2 *Material*

We used the material and the instructions of Gawronski and colleagues (2015). More specifically, the CSs were five geometrical shapes that were paired with 16 positive and 16 negative USs that were pictures from the International Affective Picture System (or IAPS; Lang et al., 1997). Among the five CSs, two were systematically paired with positive USs, two were paired with negative USs and one was presented alone.

2.1.3 *Procedure*

The procedure for Experiment 1 and 2 is displayed in Figure 1. Participants were randomly assigned to one of the three experimental conditions: the "response measured once" (or RMO) condition, the "response measured twice" (or RMT) condition and the "response measured continuously" (or RMC condition). In all three conditions, participants first provided socio-demographic data. They were then exposed to the acquisition phase. In this phase, two CSs were paired with positive USs and two CSs were paired with negative USs. Each CS was paired eight times with a different US of a given category (either positive or negative). An additional CS was also presented eight times but was not paired with any US. The neutral CS allowed having a baseline for each participant. The CSs were randomly assigned to the within-participant US valence conditions (positive, neutral, negative). The pairing sequence was as follows: a fixation cross (250ms) was followed by the CS (1000ms), which was replaced by the US (1000ms), both of which appeared in the center of the screen. The inter-trial time interval was of 1000ms.

Next to the acquisition phase, participants from the RMT and RMC conditions rated the five CSs on an evaluative scale from 1 (very unpleasant) to 7 (very pleasant). The participants assigned to the RMO condition directly proceeded to the extinction phase, without rating the CSs. In the extinction phase, each CS appeared eight more times, but this time not followed by a US. For the RMO and RMT conditions, the sequence was similar in every point to the pairing sequence in the acquisition phase but one: a blank screen replaced the presentation of the US. For the RMC condition, participants had to rate on the same evaluative scale the CS after each trial during the extinction phase. The sequence for the RMC condition was the following: a fixation cross (250ms) was followed by the CS presentation for 1000ms. A blank screen was then presented for 1000ms before the evaluative scale appeared to rate the CS. The scale was displayed until a response was given. Importantly, the CS was not re-presented on the screen with the evaluative scale. After the extinction phase, participants in all conditions rated the CSs on the same evaluative scale. They were then thanked and debriefed.

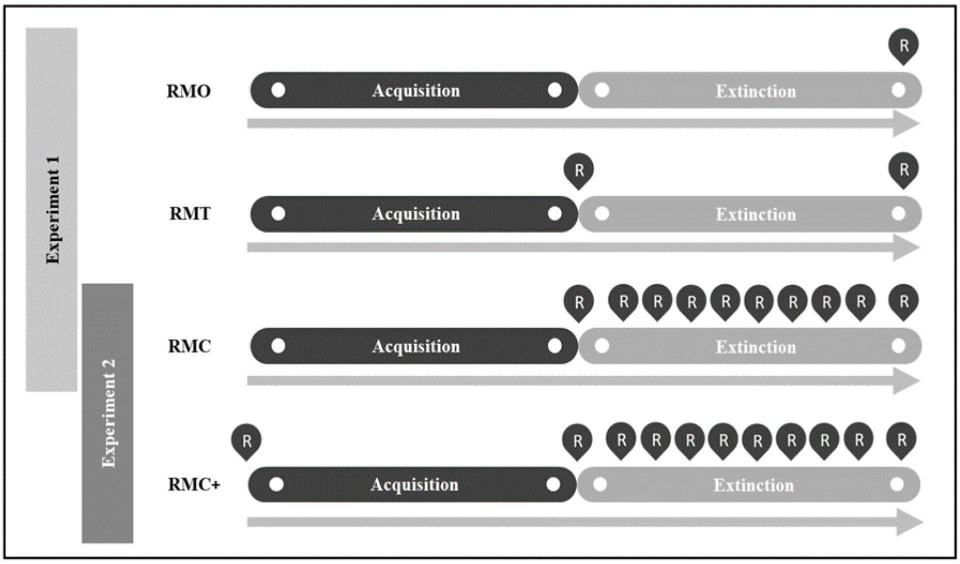


Figure 1. Experimental conditions for Experiment 1 and 2. RMO = "response measured once" condition. RMT = "response measured twice" condition. RMC = "response measured continuously" condition. RMC+ = "response measured continuously plus ratings at pre-acquisition" condition. R = Evaluative response production.

2.2 Results

2.2.1 *Analytical strategy*

For each participant, the CS ratings were averaged by CS type (positive, neutral or negative) and time of evaluation. For the main analyses, the CS ratings were baseline-corrected by subtracting the mean evaluations of both valent CSs by the mean evaluations of the neutral CSs (see Gawronski et al., 2015, for similar strategy). The data were analyzed using the 'aov' and 'anovaBF' functions in R (R Development Core Team, 2017; from the R package 'BayesFactor', Morey et al., 2015) for the frequentist and Bayesian analyses respectively. We report the Bayes factors associated with the models comparison made in the frequentist analyses. An augmented model containing the tested factor was compared to a constraint model that did not contain the tested factor. The Bayes factors in favor of the alternative hypothesis (or BF10) are presented when the conventional p-value of .05 is encountered. The Bayes factors in favor of the null hypothesis (or BF01) are reported when the p-value is above this threshold.

2.2.2 *Post-acquisition CS ratings*

The baseline-corrected CS ratings obtained during the post-acquisition phase (in the two of three experimental conditions that contained this phase) were submitted to a repeated measure ANOVA with the US valence as within-subject factor. The main effect of US valence was observed $F(1, 178) = 37.73$, $p < .001$, $\eta^2_p = .17$, BF10 > 1000 , indicating a typical EC effect: the positive CSs (M = .38, SD = 2.01) were evaluated more positively than the negative CSs (M = -.56, SD = 1.78).

2.2.3 *Post-extinction CS ratings*

The baseline-corrected CS ratings obtained during the post-extinction phase were submitted to a 2 (US valence) by 3 (Experimental condition) ANOVA, with repeated measures on the first factor. This analysis revealed a main effect of US valence $F(1, 257) = 34.26$, $p < .001$, $\eta^2 = .114$, BF10 > 1000, meaning that the EC effect was still observed after the extinction procedure: the positive CSs (M = .23, SD = 1.84) were evaluated more positively than the negative CSs (M = -.44, SD = 1.92). No main effect of condition was observed $F(2, 257) =.89$, $p = .82$, $\eta^2_p = .001$, BF01 = 8.71. The average CS ratings from the RMO (M = .08, SD = 2.03), the RMT (M = -.27, SD = 1.86) and the RMC (M = -.13, SD = 1.84) were not different from each other. The crucial interaction between US valence and the experimental condition was not observed (see Figure 2), $F(2, 257) = 1.67$, $p = .19$, $\eta^2_p = .005$, BF01 = 5.26.

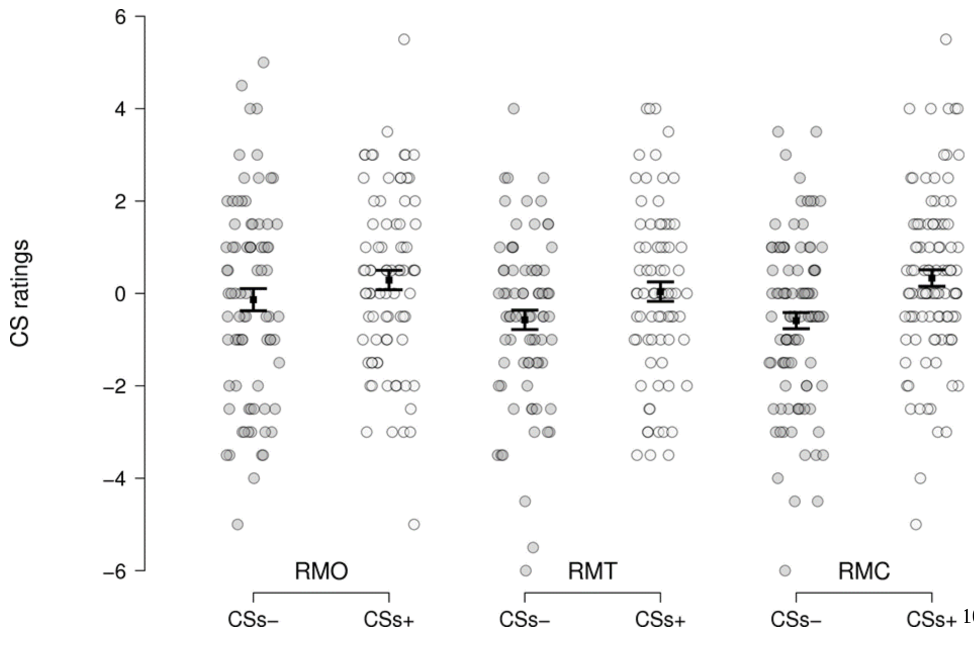


Figure 2. CS ratings represented separately for baseline-corrected positive and negative CSs in the three extinction conditions. Dots represent individual responses. The filled squares represent the observed means. Error bars represent s standard errors of means. CSs- = negatively conditioned stimuli. CSs+ = positively conditioned stimuli. RMO = “response measured once” condition. RMT = “response measured twice” condition. RMC = “response measured continuously” condition. After the extinction phase, the EC effect remained in all conditions.

2.2.4 *Comparison between post-acquisition and post-extinction CS ratings*

The baseline-corrected CS ratings were submitted to a 2 (US valence) by 3 (Experimental condition) by 2 (time) ANOVA, with repeated measures on the first and the third factors. The analysis revealed a main effect of US valence $F(1, 531) = 76.79$, $p < .001$, $\eta^2_p = .125$, BF10 > 1000, with positive CSs (M = .29, SD = 1.92) evaluated more positively than negative CSs (M = -.57, SD = 1.78). No main effect of time or condition was observed (Fs < 1). The US valence by condition interaction was not observed either (F < 1). A weak interaction between condition and time was observed $F(1, 531) = 3.95$, $p = .047$, $\eta^2_p = .006$, BF10 = 0.65, with the difference in ratings between post-extinction and post-acquisition phases being more negative in the RMT (M = -.32, SD = 2.09) than RMC (M = .07, SD = 1.83) condition. Importantly, neither the US valence by time interaction, $F(1, 531) = 0.56$, $p = .45$, $\eta^2_p = .001$, BF01 = 6.37, nor the US valence by time by condition interaction (see Figure 3), $F(1, 531) = 1.76$, $p = .186$, $\eta^2_p = .001$, BF01 = 2.80, were observed, indicating that no extinction effect was found whatsoever.

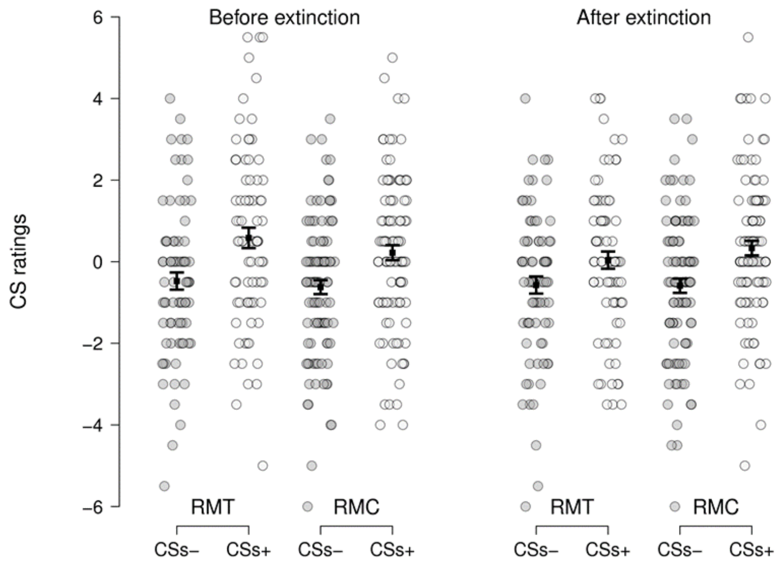


Figure 3. CS ratings represented separately for baseline-corrected positive and negative CSs, the different extinction conditions and the time of measurement (i.e., before and after the extinction phase). Dots represent individual responses. The filled squares represent the observed means. Error bars represent standard errors of means. CSs- = negative conditioned stimuli. CSs+ = positive conditioned stimuli. RMT = "response measured twice" condition. RMC = "response measured continuously" condition. EC was not different before and after the extinction phase and this was true in all experimental conditions.

2.2.5 ***Online ratings during the extinction phase***

CSs ratings obtained during the extinction phase of the RMC condition were submitted to a repeated measure ANOVA, with US valence and time of measurement (from the first to the eighth trial) as within-subjects factors. Importantly, time of measurement was coded and entered as a linear factor. We did not make the baseline-correction as before but rather analyzed the CSs of the three valent categories (positive, neutral and negative) separately. The analysis revealed a main effect of US valence, $F(2, 2341) = 170.39$, $p < .001$, $\eta^2_p = .126$, BF10 > 1000, reflecting a general EC effect (CSs- : M = 3.27, SD = 1.21; CSs neutral : M = 3.92, SD = 1.52; CSs+ : M = 4.22, SD = 1.34). We further observed a main effect of time, $F(1, 2341) = 11.61$, $p < .001$, B = - .05, $\eta^2_p = .005$, BF10 = 1.21, meaning that CSs were overall evaluated more negatively over the completion of the extinction phase. This effect is however small and inconclusive as highlighted by the Bayesian analysis. Finally, there was no interaction between US valence and time, $F(2, 2341) = 0.90$, $p = .406$, $\eta^2_p = .000$, BF01 = 63.57. The absence of such interaction suggests that EC did not extinguish (see Figure 4).

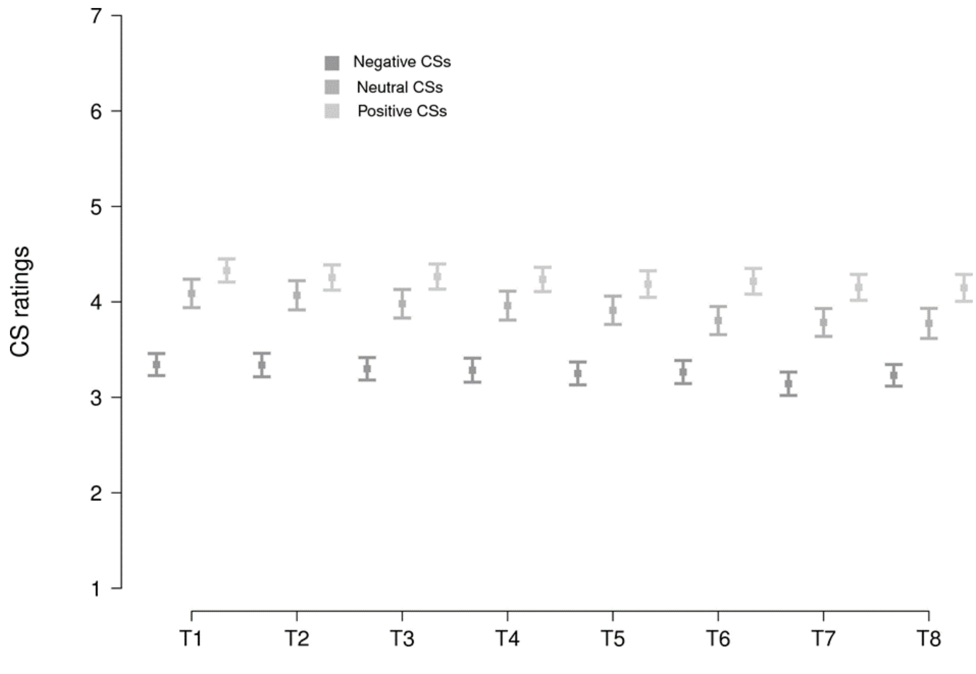


Figure 4. Online CS ratings represented separately for positive, neutral and negative CSs across eight times of measurement in the extinction phase. The filled squares represent the observed means. Error bars represent the standard errors of the means. EC did not decrease over the extinction phase in the RMC condition.

2.3 Discussion

Experiment 1 tested the extinction of EC by comparing three experimental conditions in which we varied the frequency of conditioned evaluative responses. We hypothesized that ensuring the continuous expression of evaluative responses over the extinction phase should facilitate the extinction of EC. We observed reliable EC. Despite a proper acquisition of evaluative responses, high statistical sensitivity, and the use of sensitive materials and procedures, EC did not extinguish. We did not replicate the influence of the presence vs. absence of post-acquisition evaluative ratings on the evaluative extinction that was found by Gawronski and colleagues (2015), which speaks to the instability of the latter effect.

As compared to the RMO and RMT conditions, in the RMC condition participants had to produce an evaluative response. This leads to a difference in time delay between extinction trials (that could arguably contribute to increasing extinction). Yet, even with this potentially facilitating factor, no extinction was observed.

3. Experiment 2

Experiment 2 was aimed at further facilitating extinction, this time by making clear to the participants that they may change their evaluation over the course of the experiment. This was done by examining extinction of EC-acquired in two conditions: the RMC condition designed in Experiment 1, and a variation of it that now also includes a pre-acquisition rating. In the latter condition, participants are given the opportunity to experience that their evaluations changed as a function of the experimental stage they were engaged in.

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the experiment. The raw data from the experiment as well as the R script created to manage, analyze and visually represent the data are publicly available on Open Science Framework (OSF link: osf.io/ky9tn).

3.1 Method

3.1.1 *Participants*

We collected data on 193 participants (M_{age} = 30.80, SD_{age} = 9.70, 87 women) on the online platform ‘Prolific academic’. Participants were compensated of 0.6 £ for their participation.

The latter sample size allows disposing of 99.8% chance to replicate a significant within-participants differential effect between post-extinction and post-acquisition CS ratings (unilateral alpha inferior or equal to .05) assuming a 'true' effect size of Cohen's d = .322 (as reported by Hofmann and colleagues, 2010). This sample size further allows to reliably observing a between-participants effect of the response frequency condition as small as. Cohen's $f^2 = .07$ = with a power of $(1- \beta) = .95$, and a two-tailed $\alpha = .05$

3.1.2 *Material and procedure*

The material and procedure were similar in every point to Experiment 1's RMC condition with one difference. Half of the participants took part in the RMC condition, whereas the other half additionally completed ratings prior to the acquisition phase (RMC+ condition)

3.2 Results

3.2.1 *Comparison between post-acquisition and post-extinction CS ratings*

Using the same analytical strategy as in Experiment 1, the baseline-corrected CS ratings were submitted to a 2 (US valence) by 2 (Experimental condition) by 2 (time) ANOVA, with repeated measures on the first and third factors. This analysis revealed a main effect of US valence, $F(1, 573) = 41.83$, $p < .001$, $\eta^2_p = .13$, BF10 > 1000, the CSs+ (M = .39, SD = 1.84) were evaluated more positively than the CSs- (-.23, SD = 1.83). The main effects of experimental condition and time were not observed (Fs < 1). The interactions between US valence and condition as well as between time and condition were also not observed (Fs < 1). Importantly, neither the US valence by time interaction, $F(1, 573) = .81$, $p = .368$, η^2_p = .000, BF01 = 8.01, not the three-way US valence by condition by time interaction were observed, $F(1, 573) = .19$, $p = .66$, $\eta^2_p = .000$, BF01 = 6.60 (see Figure 5). Consistently with the first experiment, no extinction was observed in this second experiment.

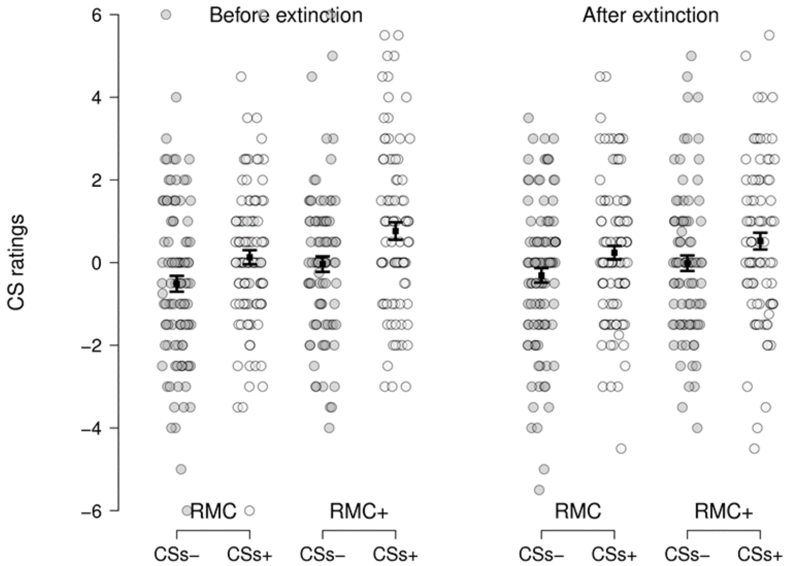


Figure 5. CS ratings represented separately for baseline-corrected positive and negative CSs, the different extinction conditions and the time of measurement (i.e., before and after the extinction phase). Dots represent individual responses. The filled squares represent the observed means. Error bars represent the standard errors of the mean. CSs- = negative conditioned stimuli. CSs+ = positive conditioned stimuli. RMC = "Response Measured Continuously" condition. RMC+ = "Response Measured Continuously + pre-acquisition ratings" condition. EC was not different before and after the extinction phase and this was true in both experimental conditions.

3.2.2 ***Online ratings during the extinction phase***

CSs ratings obtained during the extinction phase were submitted to a repeated measure ANOVA, with US valence and time of measurement (from the first to the eighth trial) as within-subject factors and the experimental condition as a between factor. Importantly, time of measurement was coded and entered as a linear factor. As for Experiment 1, we did not apply the baseline-correction but rather analyzed the CSs of the three valent categories (positive, neutral and negative) separately.

The analysis revealed a main effect of US valence, $F(2, 4411) = 76.97$, $p < .001$, $\eta^2_p = .034$, BF10 > 1000, reflecting EC (CSs- : M = 3.78, SD = 1.47; CSs neutral : M = 3.86, SD = 1.67; CSs+ : M = 4.24, SD = 1.39). No main effect of time was observed, $F(1, 4411) = 2.16$, $p = .142$, $\eta^2_p = .000$, BF01 = 3.48. We observed neither a main effect of condition nor a time by condition interaction (Fs < 1). Interestingly, we observed a weak interaction between US valence and condition, $F(2, 4429) = 6.22$, $p = .002$, $\eta^2_p = .002$, BF10 = 2.20. EC(Positive CSs – neutral CSs) – (negative CSs – neutral CSs) was stronger for RMC (M = .55, SD = 1.45) than for RMC+ (M = .39, SD = 1.34). Finally, there was no interaction between US valence and time, $F(2, 4429) = .235$, $p = .79$, $\eta^2_p = .000$, BF01= 86.65, nor an interaction between US valence, time and condition, $F(2, 4429) = .226$, $p = .80$, $\eta^2_p = .000$, BF01= 93.98. The absence of the latter interactions suggests that EC did not extinguish (see Figure 6).

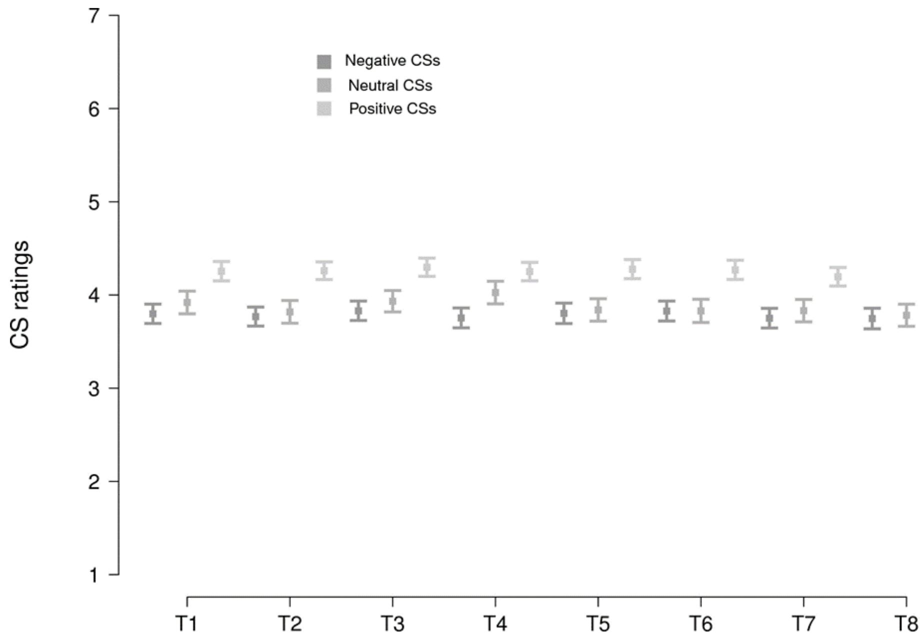


Figure 6. Online CS ratings represented separately for positive, neutral and negative CSs at eight times of measurement during the extinction phase. The filled squares represent the observed means. Error bars represent standard errors of means. EC did not decrease over the extinction phase.

3.3 Discussion

Fully consistent with Experiment 1, and so again inconsistent with the response theory of extinction, we observed no extinction of EC despite the continuous production of evaluative responses during extinction. Of interest, we observed that requesting pre-acquisition CS ratings increased (although to a weak and low reliable degree) EC. This is consistent with Gast and Rothermund's (2011) argument that pre-ratings can activate a valence-focus processing goal that, in turn, increases EC.

4. General Discussion

Drawing on the response theory of extinction (Rescorla, 1993, 1997), we predicted that the expression of CRs during the extinction procedure facilitates EC extinction. We explained in the introduction that this theory was largely consistent with available evidence on EC extinction, and that it provides a distinct (i.e., representational rather than judgmental) account of extinction effects.

To test this hypothesis, we manipulated the frequency of unreinforced conditioned responses during extinction training. In the first experiment, evaluative ratings were collected only after the extinction procedure, after both conditioning and extinction procedures, or additionally after each extinction trial. In none of these conditions did we observe an extinction of EC. In a second experiment, we compared two continuous response expression conditions, one of which additionally involved providing an evaluative rating prior to acquisition. Neither of these two continuous response expression conditions produced extinction.

The present research suggests that the presence of CS evaluations before the extinction procedure or throughout the whole extinction procedure does not suffice to elicit EC extinction. One may therefore question the notion that (1) providing CS evaluations before the extinction is sufficient for participants to form a momentary rather than an integrative CS evaluation (Gawronski et al., 2015; Lipp & Purkis, 2006), and (2) that, similarly to CC, extinction of EC depends of the formation of inhibitory link between the CS and the CR.

In Hofmann and colleagues' (2010) meta-analysis, the estimated effect size of extinction (i.e., the within-participants differential between post-extinction and post-acquisition CS ratings) was about Cohen's d = .322. The authors argued that previous experiments examining evaluative extinction were typically underpowered and might therefore have failed to observe a modest but reliable extinction of EC. The absence of extinction in the two present experiments challenges this general conclusion. As a matter of fact, the sample size of the present two studies allowed to reliably observe a within-participants differential effect of the post-extinction and post-acquisition CS ratings as small as Cohen's $d = .22$ (Exp. 1) and $d = .21$ (Exp. 2), with a power of $(1- \beta) = .90$, and a one-tailed $\alpha = .05$. This corresponds to about 1% of explained variance.

One may also note that, in the extinction studies reviewed by Hofmann et al. (2010), the post-acquisition EC was found to be larger than the average EC effect in standard EC studies. Hence, it may be argued that extinction only applies to large EC-acquired responses. Contrary to this view, the post-acquisition EC effect observed in the present experiments was as large as that observed in the extinction studies reviewed in the meta-analysis.

Finally, the present research, similar to those reviewed in the meta-analysis, relied on direct evaluative ratings. Overall, the present findings cast serious doubts on the meta-analytical conclusion that “Hence, EC is sensitive to extinction.” (Hofmann et al. 2010, p. 411), although we such effect could be induced through specific evaluative instructions implemented at the post-extinction training stage of the experiment (Aust et al., in press).

The absence of EC extinction is sometimes seen as a critical support for associative accounts of EC over propositional accounts (for a recent discussion, see Gawronski & Bodenhausen, 2018). This is because associative learning is often described as dependent on the CS-US contiguities (i.e., the absolute number of pairings), not on the CS-US contingencies (i.e., the proportion of pairings). It should be noted, however, that associative accounts point to the role of “regularities” in the environment (Gawronski, Brannon, & Bodenhausen, 2017, p. 106), implying that EC effects should be sensitive to extinction procedures. Clearly, if the CS is not statistically contingent on the US appearance, there is little regularity to be observed in the CS-US associations. This is further attested by the reference of dual-learning theorists to Hebbian models (e.g., in the APE model), the latter of which rely on statistical contingency assumptions (e.g., Caporale & Dan, 2008).

Turning to propositional learning accounts, they may equally account for the presence or absence of extinction effects depending on the underlying proposition. Extinction effects would be predicted under a “CS predicts US” propositional rule but not under a “CS co-occurs with US” rule (see De Houwer, 2018 for a recent discussion). Hence, just as for associative models, propositional models should be further constrained with regards to this question.

One could argue that ensuring the repeated expression of evaluative responses encouraged participants to be consistent, and so to report the same rating across the different stages of the experiment. This “need for consistency” may have hid a true extinction effect promoted by the repeated evaluations. In contrast to this interpretation, however, it is equally conceivable that repeatedly requesting an evaluation was actually interpreted by the participants as an invitation to modify them (if not, why asking the same question over and over again?) In other words, it is unclear why participants would have inferred that consistency rather than adaptability was desirable in the present experiments. This is particularly true in the second experiment, where we encouraged participants to figure out that their evaluation of the CSs could be changed (i.e., in the condition involving a pre-acquisition rating). More critically, whatever processes may have contributed to the absence of extinction in all of the conditions observed here, the current evidence still speaks to a functional difference between EC and CC. Whereas physiological responses acquired in a CC procedure extinguish through unreinforced response production, the extinction of evaluative responses acquired in an EC procedure does not reveal the latter sensitivity, at least *for the category of responses examined here*.

The latter italicization is important. Admittedly, future research may find evidence for EC extinction (with or without repeated response production at extinction) in paradigms that use different response formats. A major difference between CC and EC studies is that the former involves monitoring the activity of the autonomic nervous system (allowing little control over the "evaluative" responses) whereas the latter typically involve direct evaluative measures (allowing considerable control on the evaluative responses). Therefore, an interesting endeavor for future research is whether the extinction of EC effects through the production of unreinforced responses more likely applies to physiological or, more generally, to indirect evaluative responses.

To our knowledge, extinction on indirect evaluative measures has not been reported yet. Addressing the impact of repeated response expression on indirect measures would be quite challenging, though. This is because indirect measures, by definition, do not require a direct evaluation of the CS. Alternatively, one may request from participants repeated direct evaluations of the CSs during extinction training (as achieved here) and subsequently measure how this request influences an indirect evaluative measure. In our view, however, this procedure is likely to further prevent extinction effects. This is because indirect evaluative measures are largely sensitive to propositional processes (Kurdi & Banaji, 2017). In other words, repeatedly expressing one's direct evaluation of the CS ("I like it very much") is likely to strengthen rather than to weaken the outcome of the indirect measure. Using physiological measures may be more promising. In particular, it would be interesting in future research to use USs that convey both physiological and evaluative responses and to compare their relative extinction as a function of continuous conditioned response production.

Generalization across different procedures is also welcome. For instance, pairing procedures that are presumably more conducive to S-R linkage (e.g., incidental and simultaneous CS-US pairings; for a recent discussion, see Corneille & Stahl, 2018) may be more sensitive to continuous response expression. A recent study that compared extinction under pairing conditions promoting S-S or S-R linkage, however, found no differences in (pseudo)-extinction effects (Gawronski et al. 2015). Alternatively, procedures that draw participants' attention away from evaluative information may facilitate extinction of EC-acquired response. Specifically, Vanaelst, Spruyt, Everaert, and De Houwer (2017) recently observed partial extinction of EC effects when participants were led to focus their attention away from evaluative information of the CSs during extinction training. Hence, whereas orienting attention on valence may facilitate the acquisition of conditioned response (Gast & Rothermund, 2011), drawing attention away from it may facilitate extinction (Vanaelst et al., 2017).

The latter finding speaks to the dependence of both EC acquisition (e.g., Corneille et al., 2009; Gast & Rothermund, 2011; Stahl et al., 2016) and EC extinction to top-down effects of processing goals. The continuous expression of evaluative responses solicited in this experiment may have contributed to orienting participants' attention to the valence dimension of the CS at extinction. Of note, however, extinction was never observed in any of the conditions (including those where such production was not requested). Later studies should investigate further the notion that actively preventing participants from paying attention to the evaluative meaning of the CSs contributes to the extinction of the evaluative response they produce.

5. Conclusion

To conclude, two experiments conducted within the framework of Rescorla's (1993, 1997) Response Theory of Extinction examined whether the frequency of response production during extinction training plays a role in EC extinction. Two experiments failed to obtain evidence for extinction despite high statistical sensitivity and presumably conducive conditions. We conclude that if EC extinguishes, it most likely does so when specific judgment instructions influencing retrieved memory content are implemented at the evaluation stage (Aust and Stahl, in press). In turn, the latter conclusion points to an important functional difference between CC and EC, and challenges associative learning assumptions held in dual-attitude research.

GENERAL DISCUSSION

In the General Discussion, I summarize and integrate the findings of the four empirical chapters. Overall, the results are largely inconsistent with an associative learning approach to EC. I also support a single-learning model that in my view accounts for most of the current evidence in the EC literature, including evidence that is typically claimed to support dual-learning models. Specifically, in this "weak" single-learning model (Lovibond & Shanks, 2002), a propositional encoding process (requiring contingency awareness) creates both a conditioned attitude and contingency memory. Importantly, the contingency memory is *not* assumed to further determine the conditioned attitude.

1. Summary of the findings

In the four empirical chapters, we examined whether EC is produced through "associative" learning. Contrary to this view, the first two chapters challenged the view that EC is efficiently established (Chapter Two) and that conditioned attitudes observed without explicit memory of the pairings are established through implicit misattribution (Chapter Three). The next two chapters challenged the view that EC is unqualified (Chapter Four) and that it passively registers regularities in the environment (Chapter Five).

1.1 When does EC occur?

In Chapter Two, we challenged the notion that an efficient associative process drives EC (or at least a portion of it), even though we relied on a paradigm claimed to be conducive to associative learning (Hütter & Sweldens, 2013; Sweldens et al., 2010). We did so in three experiments by depleting participants' attentional resources during their exposure to the CS-US pairings. We used an auditory two-back task to prevent a visual interference with the perception of the pairings. Reducing the attentional resources systematically reduced both conditioned attitudes and (explicit) contingency memory. We furthermore found no evidence for attitudes acquired without explicit memory (i.e., the *a*-parameter was non-significant). Interestingly, in the second experiment, the resource depletion strongly reduced, though not entirely, EC. This likely resulted from an insufficient depletion manipulation, as the load manipulation also did not entirely prevent the explicit memory of the pairings.

These experiments add to previous evidence that sufficient attentional resources are necessary for the establishment of an EC effect and question the idea that specific CS-US pairing settings better allow for an efficient associative learning of attitudes (Hütter & Sweldens, 2013; Jones et al., 2009; Sweldens et al., 2010).

One may argue that the auditory two-back task interfered with the very visual processing of the pairings. As such, this "induced blindness" might have prevented associative learning. I believe this was unlikely. Using a similar two-back task, Pleyers and colleagues (2009) observed in a follow-up study a high level of recognition accuracy when participants were asked to select which stimuli were actually presented in the conditioning procedure. Recognition accuracy was however not compared to the control condition and the sample size was rather small (N = 13). This question should therefore be addressed further in future research.

In Chapter Three, we gave associative learning an even better chance to operate in EC. In two experiments, we manipulated the affective evocativeness of the US, which is seen as a critical enabling factor for an Implicit Misattribution Process to come into play (Jones et al., 2009). The IMP is to date the associative account of EC that has received the strongest empirical support (Sweldens et al., 2014). The US evocativeness manipulation allowed opposing predictions from associative and propositional accounts. Whereas the IMP (Jones et al., 2010) predicts a stronger unaware EC with low US evocativeness, the propositional approach predicts a stronger aware EC with high US evocativeness. We obtained surprising results. This time, we did observe conditioned attitude without explicit CS-US contingency memory, but one that was not dependent on US evocativeness. Although one cannot exclude the possibility that an associative process played a role in the establishment of EC in these experiments, one must reject the idea that this encoding process (if any) reflected IMP (Jones et al., 2009, 2010).

As a further insight, the experiments reported in the first two chapters support the validity of the MPT approach in that the parameters obtained by this approach were independent from each other. Reducing the memory parameter (by depleting attentional resources) or increasing the memory parameter (with higher evocativeness) did not impact the attitude parameter. These parameters might therefore reflect qualitatively rather than quantitatively distinct constructs.

Overall, based on the evidence collected here, I believe that it is fairer to see the attitude parameter as reflecting conditioned attitudes in the absence of explicit memory of the pairings, rather than reflecting weak memory of the pairings. I am however not in favor of the idea that the attitude parameter reflects conditioned attitudes that were *acquired* without awareness. In particular, Hütter and De Houwer (2017) observed an attitude parameter in instructed-based EC (i.e., in the absence of actual exposure to CS-US pairings), and that parameter further varied as a function of a manipulation implemented at retrieval (Hütter & De Houwer, 2017). As I will argue later, I believe that the awareness of the pairings is necessary for the establishment of EC, but contingency memory might not be necessary for the maintenance of EC.

1.2 What is learned through EC?

In Chapter Four, we tested whether the CS-US linkage that is created through EC leads to reciprocal or non-reciprocal evaluative changes. According to contemporary associative learning accounts of EC, CS-US associations are unqualified and should therefore lead to reciprocal evaluative changes. In contrast, a CS-US proposition may contain not only the strength of the relation but also its structure (Lagnado et al., 2007; Shanks, 2007).

In three experiments, we measured the evaluations of both CSs and USs and tested how they changed as a result of their pairing. In the first experiment, we observed that both CSs and USs evaluations were impacted by the CS-US pairings. Whereas CSs increased in valence intensity (in the direction of their paired USs), USs decreased in valence intensity (necessarily in the direction of their paired CSs). In two additional experiments, we observed that this “US devaluation effect” was likely to be driven by an affective habituation effect. It was due to repeated exposure to the USs rather than to exposure to the CS-US pairings *per se*. Therefore, we concluded that the CS-US linkage stored in memory is qualified by directionality, which seems more consistent with a propositional than with an associative approach to attitude learning. Of interest, these studies do not only suggest that the evaluative transfer in EC is non-reciprocal, but also that, by default, it goes from the US to the CS. Future research should consider whether this default can be propositionally influenced.

In Chapter Five, we tested whether the extinction of EC is observed under high statistical sensitivity conditions, and whether it depends on the production of conditioned responses during extinction training (e.g., Rescorla, 1997). This test not only provides a functional comparison of EC with Classical Conditioning; it further informs on how the CS-US link may be represented in memory. More specifically, the sensitivity or resistance of EC to extinction speaks to the potential formation of inhibitory associations between the CS and the CR (Rescorla 1994, 1997). Contrary to the latter view, and more generally to the view that associative learning passively registers regularities in the CS-US pairings, in two experiments, and across five experimental conditions varying in the frequency of the production of conditioned responses, we did not observe any evidence for an extinction of EC.

2. How can we model Evaluative Conditioning?

2.1 Dual- versus single-learning approach

As discussed in the General Introduction, there are two main categories of models of EC: dual-learning models and single-learning models. Dual-learning models posit that, in addition to propositional processes, associative processes can lead to the establishment of

EC. We also saw that "associative" learning can be defined in terms of operating conditions and operating principles of EC.

In the various studies conducted in our doctoral research (some of which only are reported in this dissertation), we obtained very little support for associative learning in EC. We observed (1) that the establishment of EC is *non*-efficient (Chapter Two), (2) that the most supported associative account of EC, the IMP, is not supported in studies making use of finer-grained MPT analyses (Chapter Three), and (3) that the CS-US linkages stored in memory is qualified by directionality (Chapter Four). We were also unable to find any indication of unaware EC in several Continuous Flash Suppression and Attentional Blink studies that were not discussed here.

To date, three lines of research are consistent with the associative approach of EC. First, EC can be obtained in incidental paradigms for which the intentionality of learning is arguably minimal (e.g., Jones et al., 2009). As discussed by Stahl and Heycke (2016), the default attentional resource allocated to the processing of the CS-US pairs in the incidental paradigms might however be sufficient for a propositional encoding of the pairings and could therefore be explained by a single-learning model. Second, the A parameter of the MPT studies was described as reflecting the operation of an associative learning process (Hütter & Sweldens, 2013). It is however now more safely described as reflecting EC without explicit memory of the pairings (Hütter & De Houwer, 2017). As discussed elsewhere (Corneille & Stahl, 2018; De Houwer, 2012; Gawronski & Walther, 2012; Sweldens et al., 2014), observing EC without contingency *memory* (i.e., explicit memory of the pairings) does not necessarily mean that EC can be established without contingency *awareness* (i.e., conscious processing of the pairings). As I discuss below, EC in the absence of contingency memory may be explained by a ("weak") single-learning model. Third, EC seems uncontrollable (Gawronski et al., 2014; Gawronski, Mitchell, & Balas, 2015; Hütter & Sweldens, 2018). The latter claim is to date empirically undisputed and therefore constitutes in my opinion the strongest (if not the only) support for the associative approach.

All in all, the evidence supporting associative learning (and, as a result, the dual-learning approach to attitude) is relatively scarce, and a "single-process learning approach to attitude acquisition may offer a parsimonious alternative to dual conceptualizations" (Corneille & Stahl, 2018, p. 70).

I now discuss two single-learning models that have been suggested in research on classical conditioning and that are relevant to EC research as well.

2.2 "strong" and "weak" single-learning models

Lovibond and Shanks (2002) formalized and reviewed empirical support to three distinct models in classical conditioning, leading from the exposure to CS-US contingency to the formation of a conditioned response and contingency memory (although they referred to it as ‘contingency awareness’). They opposed a dual-learning model to two distinct single-learning models: a “strong” and a “weak” single-learning model:

- The dual-learning model posits that a propositional process leads to contingency memory whereas an independent associative process leads to the conditioned response.
- The “strong” single-learning model posits that there is a single – propositional – learning process that leads to contingency memory, which then gives rise to the conditioned response.
- Finally, the “weak” single-learning model posits that the propositional encoding process independently leads to the formation of both contingency memory and the conditioned response. As such, contingency memory has a “weaker” role in EC. This model “could be classed as an epiphenomenalist model because awareness is granted no causal role in eliciting the CR.” (Lovibond and Shanks, 2002, p. 4).

The critical difference between these three models concerns the role of contingency memory in the production of the conditioned response. These models all predict that, in many situations, both contingency memory and the conditioned response will be observed. They however differ in predicting their correspondence. The “strong” single-learning model predicts a very close correspondence between contingency memory and the conditioned response because the latter is supposed to depend on the former. It assumes that the presentation of the CS triggers the retrieval of the CS-US contingency, which in turn elicits the conditioned response. In contrast, the dual-learning model does not predict such correspondence. Because they arise from independent learning systems, the occurrence of contingency memory and the occurrence of the conditioned response do not inform on the occurrence of each other. One can therefore expect a conditioned response without contingency memory and *vice-versa*. Finally, the “weak” single-learning model stands between the two. Contingency memory and the conditioned response may be observed independently of each other, but they should usually stand in close correspondence because they come from the same learning system. Once stored in memory, the conditioned response and the contingency memory do not relate to each other anymore. As a result, reducing contingency memory should not reduce the conditioned response.

Although EC should not be assimilated to classical conditioning, single-learning models are clearly relevant to EC. Whereas the “strong” single-process model has the strictest predictions, the “weak” single-process model can account for dissociations between conditioned attitudes and contingency memory, but without postulating the existence of two independent learning systems.

2.3 Deciding between the models

By formally distinguishing “contingency awareness” from “contingency memory”, one might specify their role in the establishment versus sustainability of conditioned attitudes. Whereas there is little convincing evidence that EC can be established without contingency awareness (Sweldens et al., 2014), it is unclear whether conditioned attitudes may be observed without contingency memory (e.g., Hütter & Sweldens, 2013; Hütter et al., 2012; Mierop et al., 2018). In terms of the three models of Lovibond and Shanks (2002), this means that whereas the dual-process model should be rejected, EC may be consistent with both the strong and the weak single process models. Because contingency memory and conditioned attitudes are both sensitive to the same encoding manipulations (Dedonder et al., 2014; Mierop et al., 2017; Pleyers et al., 2009; Stahl et al., 2016), it is reasonable to assume that they originate from a single learning system. It is however still unclear whether contingency memory should be considered a necessary condition for the sustained production of a conditioned evaluation. This distinction has much to do with the description of an attitude as either an enduring disposition (e.g., Fazio, 2007; Visser & Mirabile, 2004) or as a judgment that is constructed whenever needed (e.g., Schwarz, 2007). If episodic memory is required for the evaluation to be made, one could conclude that attitudes are constructed in the situation, based on currently accessible information.

Deciding between the weak and the strong single process models implies to test the correspondence between contingency memory and the conditioned evaluation. More specifically, in contrast to the weak single process model, the strong single process model predicts that the conditioned evaluations are formed conditionally to the retrieval of the CS-US pairings (see Figure 3). In this section, I review some work that addressed the causal role of contingency memory on conditioned evaluations.

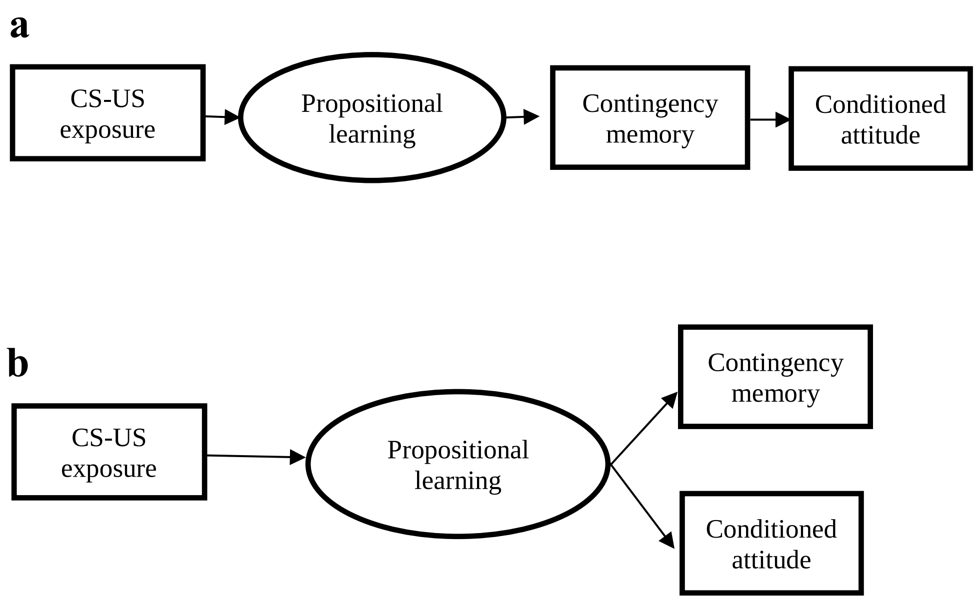


Figure 3. Single-learning approaches to EC. The CS-US exposure leads to the development of both contingency memory and conditioned attitudes through a propositional learning process. In the “strong” single-learning process (Figure 3a), contingency memory determines the conditioned attitude. In the “weak” single-learning process (Figure 3b), contingency memory and conditioned attitudes are not necessary causally related to each other.

As we saw earlier, when relying on an item-based analysis of contingency memory, EC is only observed for pairs that can be recollected after the conditioning procedure (Pleyers et al., 2007). More specifically, the US valence rather than the US identity paired with the CS has to be retrieved (Stahl et al., 2009). Interestingly, a (smaller) reversed EC is observed when the incorrect US valence is reported. This might suggest that the CS evaluation is based on the retrieval of the (occasionally reversed) memory content, but also that memory can be based on (occasionally reversed) CS evaluation. Furthermore, when memory measures are taken temporally close to the conditioning phase, the relation between EC and contingency memory could either indicate that contingency awareness is required at learning or that the recollection of the pairings is required at the CS evaluation time. As such, these experiments are not sufficient to decide between the weak and the strong single-process models.

Understanding the role of memory in EC maintenance requires to manipulating the memory content. Most of the experiments looking at the effect of a time delay on EC revealed decreased memory but stable CS evaluations over time (Baeyens et al., 1992; Förderer & Unkelbach, 2013; Grossman & Till, 1998; Hütter et al., 2012; Kerkhof et al., 2009; Walther et al., 2009). For instance, Grossman and Till (1998) examined conditioned attitudes (1) immediately, (2) one week, and (3) or three weeks after conditioning. They observed that EC remained stable over this three-weeks period, and that this stability could not be explained by measurement effects. This demonstration was carried out by relying on both a within-subjects design (Experiment 1) and a between-subjects design (Experiment 2), where participants rated the stimuli only once, either immediately after the conditioning procedure, after one week, and/or after three weeks.

Maybe more relevant to the present discussion is the relation between EC and contingency memory. Gast, De Houwer, and De Schryver (2012) collected measures on CS evaluations and US identity memory in two sessions separated by ten days. They observed that EC was found at the second session and that it was better predicted by memory observed at the second than first session. Going one step further, Förderer and Unkelbach (2013) collected CS evaluations as well as US valence and US identity measures both immediately after the conditioning procedure and seven days after. They observed that EC remained stable over time, yet US identity memory decreased over time and was not necessary for EC. However, CS-US valence memory was stable and predicted EC at the time of ratings. Overall, EC seems to be stable over time and it is accompanied by US valence memory. Once again, this correspondence can be explained by a strong (i.e., the retrieval of the pairings is necessary to form an evaluation) and a weak (i.e., the evaluation serves as a cue for the memory judgment) single-process approach.

The PD approach developed by Hütter and colleagues was specifically aimed at disentangling the contribution of explicit US valence memory and the affect-as-heuristic in memory judgments. In their fifth experiment, they measured both CS evaluations and US valence memory at two different sessions separated by 24 hours. They observed that, again, EC is stable over time. Of interest, the M-parameter decreased over time, but not the a-parameter. This could be interpreted as evidence that conditioned attitudes are maintained over time even when contingency memory is lost.

I now review a series of EC experiments that aimed at actively manipulating contingency memory. Two experimental studies manipulated CS-US contingency memory after the encoding of the CS-US pairings through “instructed forgetting” (Gast & Kattner, 2016; Molet et al., 2016). Both showed that explicitly instructing participants to forget the CS-US pairs reduced the magnitude of EC. As Richter and Gast (2017) pointed out, however, the latter studies should be “interpreted with some caution, because their design is either correlational or relies on instructions which might increase the possibility of an influence of demand compliance” (Richter & Gast, 2017, p. 2).

In order to overcome the potential impact of demand compliance, Richter and Gast (2017) manipulated time delay between each CS-US pairing with the idea that the temporal distribution of event repetition influences memory of the events (see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Donovan & Radosevich, 1999; Janiszewski, Noel, & Sawyer, 2003, for reviews). The distributed pairing procedure (i.e., each pairing was presented twice for four seconds) led to stronger contingency memory and conditioned attitudes than the massed pairing procedure (i.e., each pairing was presented once for eight seconds). As acknowledged by the authors, however, the number of CS-US co-occurrences was confounded with their time distribution manipulation. The number of co-occurrences affects the CS-US contiguity, which is known to influence EC as predicted by both associative (Baeyens et al., 1992; Jones et al., 2009; Martin & Levey, 1978) and propositional (De Houwer, 2009) accounts of EC.

In a more stringent test of the role of contingency memory on EC, the authors also tested the comparison between the distributed pairings separated for eight seconds and the distributed pairings separated for 100 seconds (maintaining the CS-US contiguity constant). They observed that the time spacing was indeed beneficial to contingency memory but did *not* influence EC. The authors interpreted the latter result as "in line with the assumption that EC is based on similar processes as memory for the pairings", yet the causal role of contingency memory on the conditioned attitudes was not supported.

In an even "stronger" single-learning model, i.e., the memory-based account of EC, Stahl and colleagues (Aust et al., *in press*; Stahl & Aust, *in press*; Stahl & Heycke, 2016) argued that the conditioned attitude and the contingency memory are not two distinct mental representations, but rather reflect two different judgments that are produced using a single set of episodic information. Recently, Aust and colleagues (*in press*) exposed their participants to two learning phases. For instance, a conditioning phase was followed by an extinction phase. They observed that participants rated the CSs as a function of the learning phase they were explicitly instructed to consider for their judgment. If they were instructed to rate the stimuli using only the information about the first – conditioning – phase, they evaluated differently the CSs+ and CSs- (suggesting an absence of extinction). If, however, they were instructed to rate the CSs using exclusively the second – extinction – phase, they did not show such differential anymore (suggesting successful extinction). In my view, although it confirms that participants are able to control the expression of their attitudes (Gawronski et al., 2014; Hütter & Sweldens, 2018), the latter evidence does not necessarily mean that participants do not have a distinct conditioned attitude that is stored in memory regardless of their contingency memory. In fact, if the conditioned attitude and the contingency memory would reflect the same set of information stored in memory, it is unlikely that any uncontrollable EC could be observed. Because controlling the expression of attitudes requires contingency memory (Balas & Gawronski, 2012), any residual uncontrolled EC would violate the assumption that EC is entirely determined by (or assimilable to) contingency memory.

All in all, whereas both contingency memory and conditioned attitudes seem to arise from the same learning process, it is to date not entirely clear whether contingency memory is necessary for the maintenance of EC. I believe it is therefore safer to favor the “weak” single-learning model. This model is more cautious in that it does not specify a (ambiguously supported) causal relationship between contingency memory and conditioned evaluations.

3. Limitations, perspective, and future research

Below I discuss several limitations in the research presented in this dissertation as well as promising avenues for future research. More specifically, I discuss the role of (1) the type of evaluative measures, (2) the conditioning paradigm, (3) the type of stimuli, and (4) the components of attitudes. I then discuss potential insights from research on brain impairment and conditioning. I end by discussing the current developments in theories of EC.

3.1 Evaluative measures

All the experiments reported here relied on direct evaluative measures (i.e., evaluative ratings). These measures imply to ask participants to directly evaluate the CSs by checking a numeric response on a single item. By using these measures, we assume that the participants are both able and willing to report their attitude. Some could argue however that these conditions are not always met. Rather, some attitudes (or aspects of attitudes) may not be accessible through introspection or may be deliberately hidden for self-presentation motives (e.g., Krosnick, Judd, & Wittenbrink, 2005; Schwarz, 2008). Alternative – indirect – evaluative measures (such as the Affective Misattribution Procedure, Payne et al., 2005; the Affective Priming Task, Fazio et al., 1995; or the Implicit Association Test, Greenwald et al., 1998) have been developed to minimize motivated response biases and to investigate components of attitudes that are, supposedly, not open to introspection. The underlying assumption of the most widely used indirect evaluative measures is that evaluative associations in memory should produce different levels of interference or facilitation in responses to evaluative stimuli on the categorization of attitude objects. Typically, differences in response times are used to infer implicit attitudes. One could argue that at least some aspects of attitudes can be acquired through associative processes in EC and that these would not be easily accessible through introspection. Consequently, it has been suggested that EC effects assessed with indirect measures may be more sensitive to associative learning processes (e.g., Gawronski & Bodenhausen, 2006; McConnell & Rydell, 2014).

Evidence is however accumulating that questions the existence of qualitative differences between attitudes inferred from direct and indirect measures. On the one hand, effects consistent with a single-learning approach were obtained on indirect measures (e.g., Bading, Stahl, & Rothermund, *unpublished*; Heycke et al., 2018; Kurdi & Banaji, 2017; Pleyers et al., 2007; Stahl et al., 2009). On the other hand, experiments reporting effects claimed to result from associative learning relied on direct evaluative measure (e.g., Jones et al., 2009; Sweldens et al., 2010). Thus, the nature of the measure (direct versus indirect) is not necessarily informative with regard to attitude learning type (propositional versus associative; see Corneille & Stahl, 2018). A safer assumption regarding the distinction between direct and indirect evaluative measures is that the former are simply more controllable than the latter (Gawronski et al., 2014, 2015a, 2015b).

3.2 Conditioning paradigms

The reported experiments should not be interpreted as revealing the general absence of associative processes in EC. Even though we did not obtain evidence for associatively driven EC in our experiments, this does not necessary mean it cannot be observed in other EC paradigms. For instance, associative processes may require to jointly (1) pair multiple USs with one CS, (2) dispose of USs that are not too salient, (3) pair a US and a CS simultaneously, (4) present the CS-US pairs incidentally. To date however, there is in my opinion no experiment that convincingly predicts and demonstrates that meeting these conditions would secure an associatively driven EC. Actually, even by manipulating directly the latter factors, the obtained results could still be explained more parsimoniously by single-learning models (e.g., Hütter & Sweldens, 2013; Mierop et al., 2018; Stahl & Heycke, 2016). More generally, as operating conditions and principles of associative EC are specified more and more precisely by associative accounts of EC (which allows a proper test of the theories), the scope of the (expected) associatively driven EC becomes consequently reduced. One should therefore question the practical and theoretical interest of dual-process models of attitude acquisition and also perhaps of implicit learning in general (Shanks, 2005).

3.3 The type of conditioned stimuli

In the present research, we used different type of CSs (objects and faces) without considering that they could influence the type of learning processes behind EC. Some researchers however argued that automatic learning of evaluations may be conditional on the type of CS used. More specifically, Ruys and Stapel (2009) argued that contingency awareness is not necessary for EC of high-novelty CSs. Contrary to this hypothesis, however, Dedonder and colleagues (2010) observed that EC of high-novelty CSs does depend on contingency awareness (see also Stahl et al., 2016) and on the availability of attentional resources.

With regard to contingency memory, we observed a dependency in the available published work[1] between the social nature of the CSs and the necessity of contingency memory for EC. Whereas all experiments reporting that EC depends on contingency memory used non-social CSs (i.e., non-words, geometrical figures, consumption products), 80% of the experiments reporting that EC does not depend on contingency memory used social CSs (i.e., human faces, fictive characters). We therefore conducted three experiments (Mierop, Teixeira, Hütter, & Corneille, *unpublished*) that compared EC for social vs. non-social CSs as well as the relative role of contingency memory. In these experiments, contingency memory moderated EC for non-social CSs but not for social CSs. The EC effect size was however overall smaller for social CSs, which could not give enough room for a moderation to be observed.

In a fourth experiment, we tested whether the particular holistic processing of the faces (e.g., Farah, 1996; Leibo, Mutch, & Poggio, 2011; Tovée, 1998) could explain the relative dependency of EC to contingency memory. As a matter of fact, one holistic account described EC as resulting from the automatic formation of a holistic CS-US representation (Levey & Martin, 1975; Martin & Levey, 1978). We therefore compared EC for upright faces to EC for reversed faces. In doing so, we were able to keep constant stimulus complexity, contrast, and luminance while manipulating the holistic processing of the faces. The reversal of the faces is indeed thought to disrupt their holistic processing (e.g., Bartlett & Searcy, 1993; Rhodes, Brake, & Atkinson, 1993). In this experiment, we observed an EC of a similar effect size for upright and reversed faces, which was furthermore moderated by contingency memory. No EC was observed for CSs when those could not be correctly related to their paired US in a later identification task.

Overall to date, there is to my best knowledge no convincing evidence that the type of CS matters with regard to the automaticity of EC.

3.4 The components of attitudes

In the research presented in this dissertation, we studied the formation of “attitudes” by testing the influence of the CS-US pairings on CS “evaluations”. One could argue that CS evaluations involve both “hot” and “cold” evaluative components. The “hot” emotional aspect would refer to emotional processes that are tightly linked to bodily activation. The “cold” emotional aspects would rather reflect propositional knowledge about emotions (e.g., Johnson & Multhaup, 1992; Leventhal & Scherer, 1987; Philippot & Schaefer, 2001). US arousal and evocativeness are probably the factors that addressed the most closely the “hot” dimension of the evaluation.

[1] This literature review was conducted in April 2015

Gawronski and Mitchell (2014) paired CSs with USs of negative or positive valence that were either highly or lowly arousing. They observed that EC was stronger when CSs were paired with highly arousing USs. They further observed that the CSs acquired the arousal of the USs and that this “arousal conditioning” was correlated with contingency memory. This is consistent with a single-process view for which the stronger arousal of the USs leads to a stronger encoding of the pairings, in turn resulting in stronger contingency memory and conditioned evaluations.

In Chapter three, we tested whether highly evocative USs may result in EC effects in the absence of awareness (or at least memory). Again, stronger evocativeness led to stronger contingency memory but was not linked with stronger (or smaller) EC for CS-US pairs that were not correctly remembered.

In sum, to date, there seems to be no empirical evidence supporting the view that a “hot” evaluation can be acquired through associative processes.

3.5 Insights from neuroscience

Multiple neuroscience theories of evaluative and fear learning share the assumption that the acquisition of emotional response can be mediated by subcortical networks, such as the amygdaloid nuclear complex, that are not under the control of higher cognitive processes (for a review see LeDoux, 2000). For instance, the Memory Systems Model (Amodio & Berg, 2018; Amodio & Ratner, 2011) assumes that a distinct memory system, physiological rooted in the amygdala and associated subcortical areas, is specifically dedicated to the acquisition, extinction and behavioral expression of emotional responses. Another popular model, the “two-pathway” model (LeDoux, 1998; Tamietto & de Gelder, 2010; Vuilleumier, 2005), assumes that a subcortical “low road” appraises the emotional significance of events unconsciously and in a way that is functionally independent of the conscious perception or the semantic categorization of the stimuli.

There is empirical evidence that the amygdaloid complex is involved in EC. The activity of the amygdala was found to be positively correlated with EC (Petrovic, Kalisch, Singer, & Dolan, 2008) and patients with brain damage involving the amygdaloid region have shown severely impaired EC (Johnsrude, Owen, White, Zhao, & Bohbot, 2000). However, the latter evidence is mixed (Coppens et al., 2006).

In six unpublished experiments, we tested whether one psychopathologic trait that is typically related with amygdala dysfunctions, i.e. psychopathic meanness (Marsh et al., 2013; Vieira et al., 2015) relates to EC (Mierop, Teixeira, Corneille, & Almeida, *unpublished*). In the first three experiments, we observed that meanness was negatively correlated with EC although it did not impact contingency memory. As such, this could be consistent with the view that evaluations and contingency memory originate from separated learning systems (dual-process model). In two additional experiments however, we did not observe such positive relation between meanness and EC when it was measured with indirect evaluative measures (i.e., the affective priming task and the implicit misattribution procedure). In a sixth experiment, we further used CSs paired with neutral USs to test for the potential asymmetry of this relation and asked participants to evaluate the USs as well as the CSs. We observed that meanness was specifically linked with lower EC of positive pairings (i.e., CSs paired with positive USs) and that it was associated with lower evaluations of the positive USs. In sum, psychopathic meanness did not impact learning *per se* (i.e., the effect of the pairings on the CS evaluations), but rather the explicit evaluation of positive stimuli irrespective of whether they had been paired or not. Because we did not observe such pattern on indirect evaluative measures, this suggests that effects occurred at judgment, not at learning.

To date, there is in my view a lack of convincing evidence that subcortical routes allow for the *automatic acquisition* of conditioned evaluations. Furthermore, it is unclear whether such subcortical networks could be *specifically* involved in affective processing or learning. Evidence is accumulating against the idea of a functionally independent subcortical route in emotional processing (e.g., Pessoa & Adolphs, 2010). In fact, the very idea of a specific network dedicated to affective processing is largely based on the finding that affective processing occurs before semantic categorization and does not require conscious perception (e.g., Hermans, Spruyt, De Houwer, & Eelen, 2003; Murphy & Zajonc, 1993; Sweeny, Grabowecky, Suzuki, & Paller, 2009). These studies however typically do not adequately verify participants' lack of awareness of the stimuli. When stricter assessments of awareness are used, evidence suggests that affective processing (1) requires conscious perception and (2) temporally follows rather than precedes semantic categorization (Lähteenmäki et al., 2015), thereby rejecting the "affective primacy" hypothesis.

Models of perception have recently started shifting from assuming distinct neuroanatomical pathways between conscious and unconscious processing to assuming different types of coordinated activity between brain regions (e.g., Dehaene et al., 2006; Lamme, 2006). Importantly, these more recent models do not assume any functional difference between affective and semantic processing.

3.6 Current developments in theorizing EC

The growing body of research on EC leads to a rapid evolution in theories of EC and, more generally, in theories of attitude formation. Because studies accumulate that are rejecting core assumptions of associative attitude learning, dual-attitude theories tend to become more complex for accommodating inconsistent data.

In a recent update of the APE model, Gawronski and Bodenhausen (2018) now argue that (1) associations and propositions should not be distinguished as mental representations and (2) the underlying associative structure is constituted of multiple layers of associations with the lower layers containing the mental representations of the stimuli and the higher layers containing the ways these are related. This revised APE model is now able to accommodate the effect of relational and validity qualifiers on resulting CS evaluations, but at the cost of removing the discriminant validity of operating principles in distinguishing the acquisition of associations and propositions.

In parallel, one of the most supported associative account of EC (i.e., the Implicit Misattribution Account; Jones et al., 2009; 2010) has been recently refined (March, Olson, & Fazio, *in press*) in a way that describes the implicit misattribution process as "relatively" rather than entirely automatic: "Inhibiting resource availability at encoding (e.g., Mierop, Hütter, & Corneille, 2017) has also been shown to interfere with EC effects; any time resources are depleted means less implicit as well as explicit processing" (March, Olson, & Fazio, *in press*, p. 26). Furthermore, it also now assumed that using a relational qualifier in an EC paradigm promotes a propositional encoding process by focusing attention on the CS. Such *ad hoc* account is hardly refutable.

The propositional models as well are becoming increasingly complex. In a recent update (De Houwer, 2018), the propositional approach now "allows for the possibility that some instances of EC have features of automaticity" (De Houwer, 2018; p. 22). The latter update allows one of the only empirical support to associative learning EC (i.e., uncontrolled EC) to be consistent with a propositional approach.

I believe that these updates have a great *descriptive* value because they allow almost all instances of EC to fit with any of the existing theoretical models. But in doing so, they are losing *predictive* value. Stated otherwise, part of these models has become irrefutable. One alternative is to turn to formalized models in which mathematical equations are describing the underlying processes as well as their sequence in EC. This new approach, recently applied to EC (i.e., Aust et al., in press; Stahl & Aust, 2018) may allow for more accurate predictions (e.g., effect sizes rather than the presence or absence of EC, learning curves). But a word of caution is certainly needed. As De Houwer (2018) stated, "because formalized models often have a protective belt of parameters that can be adjusted to rescue the model in the face of falsification (Lakatos, 1970), in reality most formalized psychological models are neither precise nor falsifiable when considering the class of all possible variants of the model." (De Houwer, 2018, p. 4).

4. Conclusion

The experiments reported in this dissertation question further the validity of associative accounts of attitude learning. As a consequence, they also question the validity of dual-models of attitudes. The efficient establishment of EC was not observed even in settings that are thought to facilitate associative encoding process. Evidence for EC in the absence of contingency memory was supported by MPT analyses. However, this effect was unstable and does not necessarily speak to learning effects. We further found that the CS-US linkage stored in memory is qualified by directionality, which is more consistent with the propositional nature of the learning process underlying EC. Finally, once established, the conditioned attitude seems to be resistant to extinction, further rejecting the idea that EC extinction reflects the passive formation of inhibitory associations. In my opinion, the current empirical evidence about EC (including the work reported in this dissertation) is best explained by a single-learning model, which posits that a propositional learning process leads to the development of both contingency memory and conditioned attitudes. The causal relation between EC and contingency memory, however, remains to be better understood. In the absence of clear-cut evidence, the most flexible “weak” single-process model should be favored. Finally, existing models of EC have all become increasingly complex, in our view at the risk of becoming less refutable.

References

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. https://doi.org/10.2307/3172667

Allen, C. T., & Madden, T. J. (1985). A Closer Look at Classical Conditioning. *Journal of Consumer Research*, *12*(3), 301. https://doi.org/10.1086/208517

Amodio, D. M., & Berg, J. J. (2018). Toward a Multiple Memory Systems Model of Attitudes and Social Cognition. *Psychological Inquiry*, *29*(1), 14-19. https://doi.org/10.1080/1047840X.2018.1435620

Amodio, D. M., & Ratner, K. G. (2011). A Memory Systems Model of Implicit Social Cognition. *Current Directions in Psychological Science*, *20*(3), 143-148. https://doi.org/10.1177/0963721411408562

Atas, A., Vermeiren, A., & Cleeremans, A. (2013). Repeating a strongly masked stimulus increases priming and awareness. *Consciousness and Cognition*, *22*(4), 1422-1430. https://doi.org/10.1016/j.concog.2013.09.011

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

Baccus, J. R., Baldwin, M. W., & Packer, D. J. (2004). Increasing Implicit Self-Esteem Through Classical Conditioning. *Psychological Science*, *15*(7), 498-502. https://doi.org/10.1111/j.0956-7976.2004.00708.x

Baeyens, F., Díaz, E., & Ruiz, G. (2005). Resistance to extinction of human evaluative conditioning using a between-subjects design. *Cognition & Emotion*, *19*(2), 245-268. https://doi.org/10.1080/02699930441000300

Baeyens, Frank, Crombez, G., De Houwer, J., & Eelen, P. (1996). No Evidence for Modulation of Evaluative Flavor–Flavor Associations in Humans. *Learning and Motivation*, *27*(2), 200-241. https://doi.org/10.1006/lmot.1996.0012

Baeyens, Frank, 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, Frank, Eelen, P., Van den Bergh, O., & Crombez, G. (1992). The content of learning in human evaluative conditioning: Acquired valence is sensitive to US-revaluation. *Learning and Motivation*, *23*(2), 200-224. https://doi.org/10.1016/0023-9690(92)90018-H

Balas, R., & Gawronski, B. (2012). On the intentional control of conditioned evaluative responses. *Learning and Motivation*, *43*(3), 89-98. https://doi.org/10.1016/j.lmot.2012.06.003

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In *Handbook of social cognition: Basic processes; Applications, Vols. 1-2, 2nd ed.* (p. 1-40). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Bartlett, J. C., & Searcy, J. (1993). Inversion and Configuration of Faces. *Cognitive Psychology*, *25*(3), 281-316. https://doi.org/10.1006/cogp.1993.1007

Bassili, J. N. (1993). Response Latency Versus Certainty as Indexes of the Strength of Voting Intentions in a Cati Survey. *Public Opinion Quarterly*, *57*(1), 54. https://doi.org/10.1086/269354

Bassili, J. N. (1995). Response Latency and the Accessibility of Voting Intentions: What Contributes to Accessibility and How it Affects Vote Choice. *Personality and Social Psychology Bulletin*, *21*(7), 686-695. https://doi.org/10.1177/0146167295217003

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review*, *6*(1), 57-86. https://doi.org/10.3758/BF03210812

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

Berkowitz, L., & Knurek, D. A. (1969). Label-mediated hostility generalization. *Journal of Personality and Social Psychology*, *13*(3), 200-206. https://doi.org/10.1037/h0028125

Bernat, E., Patrick, C. J., Benning, S. D., & Tellegen, A. (2006). Effects of picture content and intensity on affective physiological response. *Psychophysiology*, *43*(1), 93-103. https://doi.org/10.1111/j.1469-8986.2006.00380.x

Bierley, C., McSweeney, F. K., & Vannieuwkerk, R. (1985). Classical Conditioning of Preferences for Stimuli. *Journal of Consumer Research*, *12*(3), 316. https://doi.org/10.1086/208518

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

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological bulletin*, *114*(1), 80.

Bouton, M. E., Todd, T. P., Vurbic, D., & Winterbauer, N. E. (2011). Renewal after the extinction of free operant behavior. *Learning & Behavior*, *39*(1), 57-67. https://doi.org/10.3758/s13420-011-0018-6

Bouton, M. E., Trask, S., & Carranza-Jasso, R. (2016). Learning to inhibit the response during instrumental (operant) extinction. *Journal of Experimental Psychology: Animal Learning and Cognition*, *42*(3), 246-258. https://doi.org/10.1037/xan0000102

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81-105. https://doi.org/10.1037/h0046016

Caporale, N., & Dan, Y. (2008). Spike Timing–Dependent Plasticity: A Hebbian Learning Rule. *Annual Review of Neuroscience*, *31*(1), 25-46. https://doi.org/10.1146/annurev.neuro.31.060407.125639

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354-380. https://doi.org/10.1037/0033-2909.132.3.354

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and Systematic Information Processing within and beyond the Persuasion Context. *Unintended Thought*. Consulté à l'adresse https://www.scholars.northwestern.edu/en/publications/heuristic-and-systematic-information-processing-within-and-beyond

Collins, D. J., & Shanks, D. R. (2002). Momentary and integrative response strategies in causal judgment. *Memory & Cognition*, *30*(7), 1138-1147. https://doi.org/10.3758/BF03194331

Colwill, R. M. (1991). Negative discriminative stimuli provide information about the identity of omitted response-contingent outcomes. *Animal Learning & Behavior*, *19*(4), 326-336. https://doi.org/10.3758/BF03197893

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

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

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

Davies, S. R., El-Deredy, W., Zandstra, E. H., & Blanchette, I. (2012). Evidence for the role of cognitive resources in flavour–flavour evaluative conditioning. *Quarterly Journal of Experimental Psychology*, *65*(12), 2297-2308. https://doi.org/10.1080/17470218.2012.701311

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

De Houwer, J. (2007). A Conceptual and Theoretical Analysis of Evaluative Conditioning. *The Spanish Journal of Psychology*, *10*(02), 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). Why the Cognitive Approach in Psychology Would Profit From a Functional Approach and Vice Versa. *Perspectives on Psychological Science*, *6*(2), 202-209. https://doi.org/10.1177/1745691611400238

De Houwer, J. (2012). Evaluative Conditioning. In N. M. Seel (Éd.), *Encyclopedia of the Sciences of Learning* (p. 1179-1181). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4419-1428-6_1031

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

De Houwer, J., Baeyens, F., Vansteenwegen, D., & Eelen, P. (2000a). 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., Baeyens, F., Vansteenwegen, D., & Eelen, P. (2000b). 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. https://doi.org/10.1037/0097-7403.26.2.237

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. https://doi.org/10.1037/0033-2909.127.6.853

Dedonder, J., Corneille, O., Bertinchamps, D., & Yzerbyt, V. (2014). Overcoming Correlational Pitfalls: Experimental Evidence Suggests That Evaluative Conditioning Occurs for Explicit But Not Implicit Encoding of CS–US Pairings. *Social Psychological and Personality Science*, *5*(2), 250-257. https://doi.org/10.1177/1948550613490969

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

Delamater, A. R. (1996). Effects of several extinction treatments upon the integrity of Pavlovian stimulus-outcome associations. *Animal Learning & Behavior, 24*(4), 437-449. https://doi.org/10.3758/BF03199015

Díaz, E., & De la Casa, L. G. (2002). Latent inhibition in human affective learning. *Emotion*, *2*(3), 242-250. https://doi.org/10.1037/1528-3542.2.3.242

Dickinson, A. (1980). *Contemporary animal learning theory* (Vol. 1). CUP Archive.

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

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don’t. *Journal of Applied Psychology*, *84*(5), 795-805. https://doi.org/10.1037/0021-9010.84.5.795

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the Nature of Prejudice: Automatic and Controlled Processes. *Journal of Experimental Social Psychology*, *33*(5), 510-540. https://doi.org/10.1006/jesp.1997.1331

Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature, 369*, 313.

Dwyer, D. M., Jarratt, F., & Dick, K. (2007). Evaluative conditioning with foods as CSs and body shapes as USs: No evidence for sex differences, extinction, or overshadowing. *Cognition and Emotion*, *21*(2), 281-299. https://doi.org/10.1080/02699930600551592

Engelhard, I. M., Leer, A., Lange, E., & Olatunji, B. O. (2014). Shaking that icky feeling: effects of extinction and counterconditioning on disgust-related evaluative learning. *Behavior therapy*, *45*(5), 708-719. https://doi.org/10.1016/j.beth.2014.04.003

Farah, M. J. (1996). Is face recognition ‘special’? Evidence from neuropsychology. *Behavioural Brain Research*, *76*(1-2), 181-189. https://doi.org/10.1016/0166-4328(95)00198-0

Fazio, R. H. (2007). Attitudes as Object–Evaluation Associations of Varying Strength. *Social Cognition*, *25*(5), 603-637. https://doi.org/10.1521/soco.2007.25.5.603

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

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*(2), 229-238. https://doi.org/10.1037/0022-3514.50.2.229

Fiedler, K., & Hütter, M. (2014). The limits of automaticity. In *Dual Processes in Social Psychology* (Guilford Press, p. 497-513). New York: J. Sherman, B. Gawronski, & Y. Trope.

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

Field, A. P. (2000). I Like It, but I'm Not Sure Why: Can Evaluative Conditioning Occur without Conscious Awareness? *Consciousness and Cognition*, *9*(1), 13-36. https://doi.org/10.1006/ccog.1999.0402

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

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: The Role of Identity Memory, Valence Memory, and Evaluative Consolidation. *Social Psychology*, *44*(6), 380-389. https://doi.org/10.1027/1864-9335/a000150

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

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., 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., & Kattner, F. (2016). Single-trial evaluative conditioning can be moderated by instructed forgetting. *Learning & Behavior*, *44*(3), 260-269. https://doi.org/10.3758/s13420-016-0210-9

Gast, A., & Rothermund, K. (2011). 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

Gawronski, B., Balas, R., & Creighton, L. A. (2014). Can the Formation of Conditioned Attitudes Be Intentionally Controlled? *Personality and Social Psychology Bulletin*, *40*(4), 419-432. https://doi.org/10.1177/0146167213513907

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

Gawronski, B., & Bodenhausen, G. V. (2009). Operating principles versus operating conditions in the distinction between associative and propositional processes.

Behavioral and Brain Sciences, *32*(02), 207. https://doi.org/10.1017/S0140525X09000958

Gawronski, B., & Bodenhausen, G. V. (2014). The associative-propositional evaluation model: Operating principles and operating conditions of evaluation. In *Dual-process theories of the social mind* (p. 188-203).

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

Gawronski, B., Brannon, S. M., & Bodenhausen, G. V. (2017). The associative-propositional duality in the representation, formation, and expression of attitudes. *Reflective and impulsive determinants of human behavior*, 103-118.

Gawronski, B., Gast, A., & De Houwer, J. (2015). Is evaluative conditioning really resistant to extinction? Evidence for changes in evaluative judgements without changes in evaluative representations. *Cognition and Emotion*, *29*(5), 816-830. https://doi.org/10.1080/02699931.2014.947919

Gawronski, B., & Mitchell, D. G. V. (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., & Strack, F. (2004). On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology*, *40*(4), 535-542. https://doi.org/10.1016/j.jesp.2003.10.005

Gawronski, B., & Walther, E. (2012). What do memory data tell us about the role of contingency awareness in evaluative conditioning? *Journal of Experimental Social Psychology*, *48*(3), 617-623. https://doi.org/10.1016/j.jesp.2012.01.002

Gawronski, B., Walther, E., & Blank, H. (2005). Cognitive consistency and the formation of interpersonal attitudes: Cognitive balance affects the encoding of social information. *Journal of Experimental Social Psychology*, *41*(6), 618-626. https://doi.org/10.1016/j.jesp.2004.10.005

Gibson, B. (2008). Can Evaluative Conditioning Change Attitudes toward Mature Brands? New Evidence from the Implicit Association Test. *Journal of Consumer Research*, *35*(1), 178-188. https://doi.org/10.1086/527341

Gorn, G. J. (1982). The Effects of Music in Advertising on Choice Behavior: A Classical Conditioning Approach. *Journal of Marketing*, *46*(1), 94. https://doi.org/10.2307/1251163

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*(1), 3-25. https://doi.org/10.1037//0033-295X.109.1.3

Greenwald, A. G., & De Houwer, J. (2017). Unconscious conditioning: Demonstration of existence and difference from conscious conditioning. *Journal of Experimental Psychology: General*, *146*(12), 1705-1721. https://doi.org/10.1037/xge0000371

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology*, *74*(6), 1464.

Grossman, R. P., & Till, B. D. (1998). The Persistence of Classically Conditioned Brand Attitudes. *Journal of Advertising*, *27*(1), 23-31. https://doi.org/10.1080/00913367.1998.10673540

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

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

Harter, S. (1993). Causes and Consequences of Low Self-Esteem in Children and Adolescents. In R. F. Baumeister (Éd.), *Self-Esteem* (p. 87-116). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4684-8956-9_5

Hawkins, D. I., Best, R. J., & Coney, K. A. (1983). *Consumer behavior: implications for marketing strategy* (Rev. ed). Plano, Tex: Business Publications.

Healy, M. J. R., & Goldstein, H. (1978). Regression to the mean. *Annals of Human Biology*, *5*(3), 277-280. https://doi.org/10.1080/03014467800002891

Hebb, D. O. (2002). *The organization of behavior: a neuropsychological theory*. Mahwah, N.J: L. Erlbaum Associates.

Hermans, D., Dirikx, T., Vansteenwegenin, D., Baeyens, F., Van den Bergh, O., & Eelen, P. (2005). Reinstatement of fear responses in human aversive conditioning. *Behaviour research and therapy*, *43*(4), 533-551. https://doi.org/10.1016/j.brat.2004.03.013

Hermans, D., Spruyt, A., De Houwer, J., & Eelen, P. (2003). Affective priming with subliminally presented pictures. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *57*(2), 97.

Heycke, T., Gehrmann, S., Haaf, J. M., & Stahl, C. (2018). Of two minds or one? A registered replication of Rydell et al. (2006). *Cognition and Emotion*, 1-20. https://doi.org/10.1080/02699931.2018.1429389

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539-1558. https://doi.org/10.1002/sim.1186

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

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*. https://doi.org/10.1037/xlm0000533

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*(1), 79-86. https://doi.org/10.1007/s00213-010-1872-1

Houben, K., Schoenmakers, T. M., & Wiers, R. W. (2010). I didn’t feel like drinking but I don’t know why: The effects of evaluative conditioning on alcohol-related attitudes, craving and behavior. *Addictive Behaviors*, *35*(12), 1161-1163. https://doi.org/10.1016/j.addbeh.2010.08.012

Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion: psychological studies of opinion change*. Greenwood Press. Consulté à l'adresse https://books.google.be/books?id=j_FoAAAAIAAJ

Hull, C. L. (1943). Principles of behavior: An introduction to behavior theory.

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

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

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

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

Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General*, *122*(2), 139-154. https://doi.org/10.1037/0096-3445.122.2.139

Janiszewski, C., Noel, H., & Sawyer, A. G. (2003). A Meta-analysis of the Spacing Effect in Verbal Learning: Implications for Research on Advertising Repetition and Consumer Memory. *Journal of Consumer Research*, *30*(1), 138-149. https://doi.org/10.1086/374692

Janiszewski, C., & Warlop, L. (1993). The Influence of Classical Conditioning Procedures on Subsequent Attention to the Conditioned Brand. *Journal of Consumer Research*, *20*(2), 171. https://doi.org/10.1086/209342

Johnson, M. K., & Multhaup, K. S. (1992). Emotion and MEM. *The handbook of emotion and memory: Research and theory*, 33-66.

Johnsrude, I. S., Owen, A. M., White, N. M., Zhao, W. V., & Bohbot, V. (2000). Impaired Preference Conditioning after Anterior Temporal Lobe Resection in Humans. *The Journal of Neuroscience*, *20*(7), 2649-2656. https://doi.org/10.1523/JNEUROSCI.20-07-02649.2000

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), 933-948. https://doi.org/10.1037/a0014747

Jones, C. R., Olson, M. A., & Fazio, R. H. (2010). Evaluative Conditioning. In *Advances in Experimental Social Psychology* (Vol. 43, p. 205-255). Elsevier. https://doi.org/10.1016/S0065-2601(10)43005-1

Kellaris, J. J., & Cox, A. D. (1989). The Effects of Background Music in Advertising: A Reassessment. *Journal of Consumer Research*, *16*(1), 113. https://doi.org/10.1086/209199

Kendrick, R. V., & Olson, M. A. (2012). When feeling right leads to being right in the reporting of implicitly-formed attitudes, or how I learned to stop worrying and trust my gut. *Journal of Experimental Social Psychology*, *48*(6), 1316-1321. https://doi.org/10.1016/j.jesp.2012.05.008

Kerkhof, I., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2009). A picture-flavour paradigm for studying complex conditioning processes in food preference learning. *Appetite*, *53*(3), 303-308. https://doi.org/10.1016/j.appet.2009.07.007

Kerkhof, I., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2011). Counterconditioning: An Effective Technique for Changing Conditioned Preferences. *Experimental Psychology*, *58*(1), 31-38. https://doi.org/10.1027/1618-3169/a000063

Kim, J., Allen, C. T., & Kardes, F. R. (1996). An Investigation of the Mediational Mechanisms Underlying Attitudinal Conditioning. *Journal of Marketing Research*, *33*(3), 318. https://doi.org/10.2307/3152128

Kim, J. C., Sweldens, S., & Hütter, M. (2015). 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/10.1177/1948550615599237

Klauer, K. C., Stahl, C., & Voss, A. (2011). Multinomial models and diffusion models. In *Cognitive methods in social psychology* (Dialnet, p. 331-354). La Rioja, Spain: A. Voss, C. Stahl, K. C. Klauer.

Kroeber-Riel, W. (1979). Activation Research: Psychobiological Approaches in Consumer Research. *Journal of Consumer Research*, *5*(4), 240. https://doi.org/10.1086/208736

Krosnick, J. A., Judd, C. M., & Wittenbrink, B. (2005). The measurement of attitudes. *The handbook of attitudes*, *21*, 76.

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review*, *118*(1), 97-109. https://doi.org/10.1037/a0020762

Kurdi, B., & Banaji, M. R. (2017). Repeated evaluative pairings and evaluative statements: How effectively do they shift implicit attitudes? *Journal of Experimental Psychology: General*, *146*(2), 194-213. https://doi.org/10.1037/xge0000239

Lagnado, D. A., Waldmann, M. R., Hagmayer, Y., & Sloman, S. A. (2007). Beyond covariation. *Causal learning: Psychology, philosophy, and computation*, 154-172.

Lähteenmäki, M., Hyönä, J., Koivisto, M., & Nummenmaa, L. (2015). Affective processing requires awareness. *Journal of Experimental Psychology: General*, *144*(2), 339-365. https://doi.org/10.1037/xge0000040

Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, *10*(11), 494-501. https://doi.org/10.1016/j.tics.2006.09.001

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention*, 39-58.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1999). International affective picture system (IAPS): Instruction manual and affective ratings. *The center for research in psychophysiology, University of Florida*.

Leary, M. R., Schreindorfer, L. S., & Haupt, A. L. (1995). The Role of Low Self-Esteem in Emotional and Behavioral Problems: Why is Low Self-Esteem Dysfunctional? *Journal of Social and Clinical Psychology*, *14*(3), 297-314. https://doi.org/10.1521/jscp.1995.14.3.297

LeDoux, J. (1998). *The emotional brain: The mysterious underpinnings of emotional life*. Simon and Schuster.

LeDoux, J. E. (2000). Emotion Circuits in the Brain. *Annual Review of Neuroscience, 23*(1), 155-184. https://doi.org/10.1146/annurev.neuro.23.1.155

Leibo, J. Z., Mutch, J., & Poggio, T. (2011). Why the brain separates face recognition from object recognition. In *Advances in neural information processing systems* (p. 711-719).

Leventhal, A. M., Martin, R. L., Seals, R. W., Tapia, E., & Rehm, L. P. (2007). Investigating the dynamics of affect: Psychological mechanisms of affective habituation to pleasurable stimuli. *Motivation and Emotion*, *31*(2), 145-157. https://doi.org/10.1007/s11031-007-9059-8

Leventhal, H., & Scherer, K. (1987). The Relationship of Emotion to Cognition: A Functional Approach to a Semantic Controversy. *Cognition & Emotion*, *1*(1), 3-28. https://doi.org/10.1080/02699938708408361

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

Lipp, O. V., Oughton, N., & LeLievre, J. (2003). Evaluative learning in human Pavlovian conditioning: Extinct, but still there? *Learning and Motivation*, *34*(3), 219-239.

Lipp, O. V., & Purkis, H. M. (2006). The effects of assessment type on verbal ratings of conditional stimulus valence and contingency judgments: Implications for the extinction of evaluative learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *32*(4), 431. https://doi.org/10.1037/0097-7403.32.4.431

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

March, D., Olson, M. A., & Fazio, R. H. (in press). The Implicit Misattribution Model of Evaluative Conditioning. *Social Psychological Bulletin*.

Marsh, A. A., Finger, E. C., Fowler, K. A., Adalio, C. J., Jurkowitz, I. T. N., Schechter, J. C., … Blair, R. J. R. (2013). Empathic responsiveness in amygdala and anterior cingulate cortex in youths with psychopathic traits. *Journal of Child Psychology and Psychiatry*, *54*(8), 900-910. https://doi.org/10.1111/jcpp.12063

Martin, I., & Levey, A. B. (1978). Evaluative conditioning. *Advances in Behaviour research and Therapy*, *1*(2), 57-101.

Matute, H., Vegas, S., & De Marez, P.-J. (2002). Flexible use of recent information in causal and predictive judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(4), 714-725. https://doi.org/10.1037/0278-7393.28.4.714

McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, Discriminatory Behavior, and Explicit Measures of Racial Attitudes. *Journal of Experimental Social Psychology*, *37*(5), 435-442. https://doi.org/10.1006/jesp.2000.1470

McConnell, A. R., & Rydell, R. J. (2014). The systems of evaluation model: A Dual-systems approach to attitudes. *Dual-process theories of the social mind.*, 204-218.

McSweeney, F. K., & Bierley, C. (1984). Recent Developments in Classical Conditioning. *Journal of Consumer Research*, *11*(2), 619. https://doi.org/10.1086/208999

Melnikoff, D. E., & Bargh, J. A. (2018). The Mythical Number Two. *Trends in Cognitive Sciences*, *22*(4), 280-293. https://doi.org/10.1016/j.tics.2018.02.001

Mierop, A., Bret, A., Yzerbyt, V., Dumas, R., & Corneille, O. (under review). Evaluative influences of CS-US pairings are unidirectional. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Mierop, A., Hütter, M., Bret, A., & Corneille, O. (unpublished). How do processing goals influence evaluative conditioning? New insights from indirect evaluative measures and multinomial processing tree modeling. *Unpublished manuscript*.

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., Hütter, M., Stahl, C., & Corneille, O. (2018). Does attitude acquisition in evaluative conditioning without explicit CS-US memory reflect implicit misattribution of affect? *Cognition and Emotion*, 1-12. https://doi.org/10.1080/02699931.2018.1435505

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*(02), 183. https://doi.org/10.1017/S0140525X09000855

Molet, M., Kosinski, T., Craddock, P., Miguez, G., Mash, L. E., & Miller, R. R. (2016). Attenuating social affective learning effects with Memory Suppression manipulations. *Acta Psychologica*, *164*, 136-143. https://doi.org/10.1016/j.actpsy.2016.01.001

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

Morey, R. D., Rouder, J. N., & Jamil, T. (2015). BayesFactor: Computation of Bayes factors for common designs. *R package version 0.9*, *9*, 2014.

Mullen, B., Futrell, D., Stairs, D., Tice, D. M., Baumeister, R. F., Dawson, K. E., … Rosenfeld, P. (1986). Newcasters' facial expressions and voting behavior of viewers: Can a smile elect a president? *Journal of Personality and Social Psychology*, *51*(2), 291-295. https://doi.org/10.1037/0022-3514.51.2.291

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: affective priming with optimal and suboptimal stimulus exposures. *Journal of personality and social psychology*, *64*(5), 723.

Newell, B. R., & Shanks, D. R. (2014). Unconscious influences on decision making: A critical review. *Behavioral and Brain Sciences*, *38*(01), 1-19. https://doi.org/10.1017/S0140525X12003214

Nord, W. R., & Peter, J. P. (1980). A Behavior Modification Perspective on Marketing. *Journal of Marketing*, *44*(2), 36. https://doi.org/10.2307/1249975

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. https://doi.org/10.1177/0146167205284004

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning.* Oxford, England: Univer. Illinois Press.

Parker, J. S., & Benson, M. J. (2004). Parent-adolescent relations and adolescent functioning: self-esteem, substance abuse, and delinquency. *Adolescence*, *39*(155), 519-530.

Pavlov, I. P. (1927). Conditional reflexes: an investigation of the physiological activity of the cerebral cortex.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*(3), 277-293. https://doi.org/10.1037/0022-3514.89.3.277

Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: from a « low road » to « many roads » of evaluating biological significance. *Nature Reviews Neuroscience, 11*, 773.

Peters, K. R., & Gawronski, B. (2011). Are We Puppets on a String? Comparing the Impact of Contingency and Validity on Implicit and Explicit Evaluations. *Personality and Social Psychology Bulletin*, *37*(4), 557-569. https://doi.org/10.1177/0146167211400423

Petrovic, P., Kalisch, R., Singer, T., & Dolan, R. J. (2008). Oxytocin Attenuates Affective Evaluations of Conditioned Faces and Amygdala Activity. *Journal of Neuroscience, 28*(26), 6607-6615. https://doi.org/10.1523/JNEUROSCI.4572-07.2008

Petty, R. E., & Cacioppo, J. T. (1986). Message elaboration versus peripheral cues. In *Communication and persuasion* (p. 141-172). Springer.

Petty, R. E., & Wegener, D. T. (1999). The elaboration likelihood model: Current status and controversies. *Dual-process theories in social psychology.*, 37-72.

Petty, R. E., Wegener, D. T., & Fabrigar, L. R. (1997). ATTITUDES AND ATTITUDE CHANGE. *Annual Review of Psychology*, *48*(1), 609-647. https://doi.org/10.1146/annurev.psych.48.1.609

Philippot, P., & Schaefer, A. (2001). Emotion and memory.

Pleyers, G., Corneille, O., Luminet, O., & Yzerbyt, V. (2007). Aware and (Dis)Liking: Item-Based Analyses Reveal That Valence Acquisition via Evaluative Conditioning Emerges Only When There Is Contingency Awareness. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, *33*(1), 130-144. https://doi.org/10.1037/0278-7393.33.1.130

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

Pratte, M. S., & Rouder, J. N. (2009). A task-difficulty artifact in subliminal priming. *Attention, Perception, & Psychophysics*, *71*(6), 1276-1283. https://doi.org/10.3758/APP.71.6.1276

R Development Core Team. (2017). *A language and environment for statistical computing.* Vienna, Austria: R foundation for Statistical Computing. Consulté à l'adresse http://www.R-project.org

Razran, G. (1954). The conditioned evocation of attitudes (cognitive conditioning?). *Journal of Experimental Psychology*, *48*(4), 278-282. https://doi.org/10.1037/h0058778

Razran, G. H. S. (1938). Transposition of relational responses and generalization of conditioned responses. *Psychological Review*, *45*(6), 532-538. https://doi.org/10.1037/h0056291

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, *43*(3), 151.

Rescorla, R. A. (1993). Inhibitory associations between S and R in extinction. *Animal Learning & Behavior*, *21*(4), 327-336. https://doi.org/10.3758/BF03197998

Rescorla, R. A. (2001). Are associative changes in acquisition and extinction negatively accelerated? *Journal of Experimental Psychology: Animal Behavior Processes*, *27*(4), 307-315. https://doi.org/10.1037/0097-7403.27.4.307

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition*, *47*(1), 25-57. https://doi.org/10.1016/0010-0277(93)90061-Y

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

Robert A. Rescorla. (1997). Response Inhibition in Extinction. *The Quarterly Journal of Experimental Psychology Section B*, *50*(3b), 238-252. https://doi.org/10.1080/713932655

Rosenberg, M., Schooler, C., Schoenbach, C., & Rosenberg, F. (1995). Global Self-Esteem and Specific Self-Esteem: Different Concepts, Different Outcomes. *American Sociological Review*, *60*(1), 141. https://doi.org/10.2307/2096350

Rozin, P. (1982). "Taste-smell confusions" and the duality of the olfactory sense. *Perception & Psychophysics*, *31*(4), 397-401. https://doi.org/10.3758/BF03202667

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., 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

Ruys, K. I., & Stapel, D. A. (2009). Learning to like or dislike by association: No need for contingency awareness. *Journal of Experimental Social Psychology*, *45*(6), 1277-1280. https://doi.org/10.1016/j.jesp.2009.06.012

Rydell, R. J., McConnell, A. R., Mackie, D. M., & Strain, L. M. (2006). Of Two Minds: Forming and Changing Valence-Inconsistent Implicit and Explicit Attitudes.

Psychological Science, *17*(11), 954-958. https://doi.org/10.1111/j.1467-9280.2006.01811.x

Schwarz, N. (2007). Attitude Construction: Evaluation in Context. *Social Cognition*, *25*(5), 638-656. https://doi.org/10.1521/soco.2007.25.5.638

Schwarz, N. (2008). Attitude measurement. *Attitudes and attitude change*, *3*, 41-60.

Shanks, D. R. (2007). Associationism and cognition: Human contingency learning at 25. *Quarterly Journal of Experimental Psychology*, *60*(3), 291-309. https://doi.org/10.1080/17470210601000581

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, *17*(03), 367-395. https://doi.org/10.1017/S0140525X00035032

Sherif, M., & Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication and attitude change.* Oxford, England: Yale Univer. 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. https://doi.org/10.1086/209236

Smith, E. R. (1996). What do connectionism and social psychology offer each other? *Journal of Personality and Social Psychology*, *70*(5), 893-912.

Smith, Eliot R., & DeCoster, J. (2000). Dual-Process Models in Social and Cognitive Psychology: Conceptual Integration and Links to Underlying Memory Systems. *Personality and Social Psychology Review*, *4*(2), 108-131. https://doi.org/10.1207/S15327957PSPR0402_01

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

Stacy, A. W., & Wiers, R. W. (2010). Implicit Cognition and Addiction: A Tool for Explaining Paradoxical Behavior. *Annual Review of Clinical Psychology*, *6*(1), 551-575. https://doi.org/10.1146/annurev.clinpsy.121208.131444

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

Sternberg, D. A., & McClelland, J. L. (2012). Two Mechanisms of Human Contingency Learning. *Psychological Science*, *23*(1), 59-68.

Strack, F., & Deutsch, R. (2004). Reflective and Impulsive Determinants of Social Behavior. *Personality and Social Psychology Review*, *8*(3), 220-247. https://doi.org/10.1207/s15327957pspr0803_1

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. https://doi.org/10.1086/209117

Sweeny, T. D., Grabowecky, M., Suzuki, S., & Paller, K. A. (2009). Long-lasting effects of subliminal affective priming from facial expressions. *Consciousness and Cognition*, *18*(4), 929-938. https://doi.org/10.1016/j.concog.2009.07.011

Sweldens, S. (2018). Putting Operating Principles (What) Before Conditions (How and When) to Improve Theorizing: S – S versus S – R Learning in Evaluative Conditioning. https://doi.org/10.13140/rg.2.2.14882.12488

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., Osselaer, V., J, 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

Tamietto, M., & de Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience*, *11*, 697.

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

Tovée, M. J. (1998). Is face processing special? *Neuron*, *21*(6), 1239-1242.

Tulving, E. (1972). Episodic and semantic memory. *Organization of memory*, *1*, 381-403.

Vanaelst, J., Spruyt, A., Everaert, T., & De Houwer, J. (2017). Extinction of likes and dislikes: effects of feature-specific attention allocation. *Cognition and Emotion*, *31*(8), 1595-1609. https://doi.org/10.1080/02699931.2016.1250724

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. https://doi.org/10.1080/02699931.2014.947919

Verwijmeren, T., Karremans, J. C., Stroebe, W., & Wigboldus, D. H. J. (2012). Goal relevance moderates evaluative conditioning effects. *Learning and Motivation*, *43*(3), 107-115. https://doi.org/10.1016/j.lmot.2012.06.002

Vieira, J. B., Ferreira-Santos, F., Almeida, P. R., Barbosa, F., Marques-Teixeira, J., & Marsh, A. A. (2015). Psychopathic traits are associated with cortical and subcortical volume alterations in healthy individuals. *Social Cognitive and Affective Neuroscience*, *10*(12), 1693-1704. https://doi.org/10.1093/scan/nsv062

Visser, P. S., & Mirabile, R. R. (2004). Attitudes in the social context: the impact of social network composition on individual-level attitude strength. *Journal of personality and social psychology*, *87*(6), 779.

Vogel, T., Hütter, M., & Gebauer, J. E. (2017). Is Evaluative Conditioning Moderated by Big Five Personality Traits? *Social Psychological and Personality Science*, 194855061774019. https://doi.org/10.1177/1948550617740193

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*(12), 585-594. https://doi.org/10.1016/j.tics.2005.10.011

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., Gawronski, B., Blank, H., & Langer, T. (2009). Changing likes and dislikes through the back door: The US-revaluation effect. *Cognition & Emotion*, *23*(5), 889-917. https://doi.org/10.1080/02699930802212423

Walther, E., Nagengast, B., & Trasselli, C. (2005). Evaluative conditioning in social psychology: Facts and speculations. *Cognition & 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*(4), 233-241. https://doi.org/10.3758/BF03206429

Weinberger, J., & Westen, D. (2008). RATS, We Should Have Used Clinton: Subliminal Priming in Political Campaigns. *Political Psychology*, *29*(5), 631-651. https://doi.org/10.1111/j.1467-9221.2008.00658.x

Wiers, R. W., & Stacy, A. W. (2006). Implicit Cognition and Addiction. *Current Directions in Psychological Science*, *15*(6), 292-296. https://doi.org/10.1111/j.1467-8721.2006.00455.x

Zanon, R., De Houwer, J., & Gast, A. (2012). Context effects in evaluative conditioning of implicit evaluations. *Learning and Motivation*, *43*(3), 155-165. https://doi.org/10.1016/j.lmot.2012.02.003

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