



On the contribution of repeated experience to evaluative learning: An approach-avoidance procedure investigation

Cédric Batailler

► To cite this version:

Cédric Batailler. On the contribution of repeated experience to evaluative learning: An approach-avoidance procedure investigation. Psychology. Université Grenoble Alpes [2020-..], 2021. English. NNT: 2021GRALS028 . tel-03632869

HAL Id: tel-03632869

<https://theses.hal.science/tel-03632869>

Submitted on 6 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : **Sciences cognitives, psychologie & neurocognition**

Arrêté ministériel : 25 mai 2016

Présentée par

Cédric BATAILLER

Thèse dirigée par **Dominique MULLER**, Univ. Grenoble Alpes
et co-encadrée par **Cécile NURRA**, Univ. Grenoble Alpes

préparée au sein du **LIP/PC2S**
dans l'**École Doctorale Ingénierie pour la Santé, la Cognition
et l' Environnement**

Contribution des expériences répétées à l'apprentissage évaluatif : Une investigation basée sur les procédures d'approche et d'évitement

Thèse soutenue publiquement le **6 Décembre 2021**,
devant le jury composé de :

Pr. Dominique MULLER

Professeur à l'Université Grenoble Alpes, Directeur de thèse

Dr. Cécile NURRA

Maitre de Conférences à l'Université Grenoble Alpes,
Co-encadrante de thèse, Invitée

Pr. Maria AUGUSTINOVA

Professeure à l'Université de Rouen Normandie, Rapporteuse

Dr. Christophe BLAISON

Maitre de Conférences HDR à l'Université de Paris, Examineur

Pr. Olivier CORNEILLE

Professeur à l'Université Catholique de Louvain, Rapporteur

Pr. Anne GAST

Professeure à l'Universität zu Köln, Examinatrice

Pr. François RIC

Professeur à l'Université de Bordeaux, Examineur, Président du jury

Pr. Annique SMEDING

Professeure à l'Université de Savoie Mont-Blanc, Invitée



Acknowledgements

Pour les lecteur·ices de ce manuscrit, les mots qui suivent seront peut-être les premiers (les seuls ?) à être lus. Pour moi, ils furent les derniers à être écrits. L'occasion de marquer cette aventure d'un point final. Tout comme pour le reste de cette thèse, pour écrire ces remerciements, j'ai fait mes recherches. Et, après des dizaines et dizaines de manuscrits consultés, force est de constater qu'il n'existe pas de recette. Il existe des alexandrins, des bullets-points, des entrées de journal intimes, des billets sobres, des blagues ; tous ayant un commun le fait d'illustrer le fait qu'une thèse n'est pas une aventure qui se vit seul·e. En l'absence de recette, c'est ce point que j'ai décidé de retenir. Si la couverture d'un tel document est souvent avare en patronymes, il est bien souvent impossible de compter sur les doigts d'une main les personnes qui contribuent, d'une manière ou d'une autre, à la réalisation de ces aventures universitaires. A travers ces quelques paragraphes, je tenterai du mieux que je peux de rendre un hommage à ces personnes qui m'ont accompagné. Sûrement maladroitement, sûrement avec des oublis. À ma manière.

En premier lieu, je tiens à adresser un merci tout particulier aux personnes qui ont accepté d'être rapporteur·es de cette thèse : Olivier Corneille—merci Olivier pour tes précieux conseils au fil des ans—et Maria Augustinova. J'associe à ces remerciements l'ensemble des personnes qui ont accepté de composer mon jury : Christophe Blaison, Anne Gast, Annique Smeding et François Ric. Pouvoir discuter avec vous des travaux que nous avons réalisé fut un immense honneur pour moi. It was a great honor to be able to discuss this work with you.

Viennent ensuite les pierres angulaires de ce travail de thèse, celle et celui qui ont su maintenir ce manuscrit (et ma personne) debout. Dominique, en ayant la chance de travailler avec toi depuis sept ans—à l'heure où j'écris ces mots—je peux mesurer à quel point la tâche de mentor est ardue, et à quel point tu excelles dans ce domaine. Réussir à insuffler (et

entretenir) la curiosité, la rigueur, l'humilité, et le goût pour l'apprentissage nécessaire à ton métier me paraît représenter un défi vertigineux. Encore plus si l'on considère que pour chaque apprenant·e, il faut adapter le dosage. Au cours de cette thèse, et j'espère que tu me pardonneras pour la métaphore, tu auras été un mixologue fantastique. Cécile, je sais que parfois (souvent ?) au cours de nos réunions, tu as pu te demander comment nous sommes passés des maths et du soi à *ça*. J'espère que tu sais que de mon côté, je peux voir où se sont glissées toutes tes remarques, conseils et autres pertinences dans ce travail de thèse. Tu es sûrement trop humble pour l'avoir remarqué, mais je serai incapable de quantifier tout ce que j'ai appris avec toi. Pour moi, ce travail de thèse, à trois, est une expérience qui me marquera durablement, et à travers laquelle je pense avoir grandi. Cela aurait été impossible sans vous.

Travailler avec vous deux, cependant, ne signifie pas que travailler avec vous deux. Cela signifie surtout travailler en équipe. Ainsi, un grand merci à mon grand frère de thèse, Nicolas Bochart, et également à ma grande sœur, Marine Rougier. Merci également à la fratrie qui m'a précédé : Marie-Pierre Fayant, Anthony Lantian. J'aurais difficilement pu tomber dans un endroit plus accueillant. Merci également à celles et ceux qui ont fait partie de cette famille élargie, à tous les étudiants et toutes les étudiantes qui ont collaboré avec nous, à celles et ceux qui ont participé aux réunions du lundi, à celles et ceux qui ont travaillé sur ce travail de thèse. Il n'aurait pas été possible sans votre aide. Un merci tout particulier à Maude et Mae, ce fut un plaisir de travailler avec vous. Merci aussi à Serena et à Alicia. Un immense merci également à Adèle *Cottinad* Cottin, pour ton soutien dans bien trop d'étapes de cette thèse, merci.

Even if a lot of this doctoral dissertation happened in France, I also had the chance to visit the Lone Star State. Among the people with whom I had the chance to learn during these doctoral years, Bertram Gawronski takes an important place. I want to thank him for having

me in Austin. It was a short stay, but one in which I learned a lot on how to actually make theories and data discuss with each other.

Si j'ai pu remercier jusqu'ici celles et ceux avec qui j'ai pu collaborer étroitement, je n'ai pas encore adressé mes remerciements à celles et ceux qui ont su tenir une structure dans laquelle j'ai pu faire ma thèse. Ainsi, je remercie l'ED ISCE, et en particulier Chris Moulin, Orélie Garzena et Anne Guérin. Ce travail n'aurait pas été possible sans vous.

Un grand merci également à Anne *Chab* Chabaud. Merci d'avoir tout su toujours faire tout fonctionner. Merci pour ta bonne humeur. Merci pour tout.

Merci à l'ensemble des grand·es de la recherche. Celles et ceux des ateliers « marguerite », du département de psychologie, et d'autour. Un merci tout particulier à Céline Baeyens, Arnaud Carré, Alan Chauvin, Aurélie Graton, Florian Delmas, Hans IJzerman, Jean-Louis Monestès, Richard Palluel-Germain, Jean-Charles Quinton, Stéphane Rousset et Annique Semding (encore). C'est en grande partie grâce à vous que j'ai développé cette curiosité pour les êtres humains et que j'ai appris qu'il fallait de la rigueur pour en faire un objet d'étude.

Durant cette aventure, en plus de côtoyer les grands, j'ai eu, à certaines occasions, eu la chance de chausser leurs bottes, de me frotter à l'enseignement. Cette expérience aura été pour moi extrêmement enrichissante, et j'y ai rencontré un nombre de personnes formidable. Merci Anahi, Antoine, Gwenaëlle, Juliette, Nassima, Natalia, Warren, Yann. Vous m'aurez vous beaucoup appris. J'ai également une pensée toute particulière pour l'étudiant et l'étudiantes (brillant·es) que j'ai vu passer de l'autre côté. Lise, Yoann, c'est auprès de vous que j'ai vu le plus les années passer. Quand de petites créatures curieuses et studieuses vous êtes devenus des collègues dont les compétences dépassent déjà pleinement les miennes, je

sais que je peux me dire que je suis fier d’avoir eu la chance de côtoyer de personnes aussi chouettes que vous. Vous irez loin, je mise quelques sandwichs-raclettes dessus.

An other huge thank you, also, for my Austinite friends, my fastest friends (Ashokkumar, 2021). Miti, Ashwini, I really feel lucky that life made us stumbled onto each other. I’m looking forward to another Bennu with you. In person. Also, I hate you for making me a cheesy person.

Rémi, il y a plusieurs cases dans lesquelles j’aurais pu te ranger, histoire de donner un peu de structure à ces remerciements. Finalement, je pense que tu es une case à toi toute seule. Tu ne te résumeras ni au D213, ni aux psychopotes, ni aux réunions du Lundi, ni aux nombre de bières que nous avons partagé. Merci pour ta franchise, ton inconditionnel soutien et d’avoir, au fil des années, fait germer en moi un semblant de culture footballistique. Pavard.

Un merci aussi à Ivane pour tout soutien, et pour m’avoir permis d’explorer la partie kleptomane de ma personnalité. Merci à l’ensemble de mes copain·ines de l’ADRIPS également, et tout particulièrement à Luc, Lucie-Anna et Camille. Merci également à Amélie (Kaeloo), Brice et Vienne. Merci également à Alex (mon role model de la recherche) et à Ladislav. Avec nos thèses, je me dis que nous avons peut-être offert au Nostrum quelques semaines d’existence supplémentaire. J’espère que nous saurons un jour si sa fermeture était due au fait qu’un double expresso y était servi dans deux tasses.

Merci également à l’ensemble des (post) doc du labo, présent et passé, du D211 au D213. Merci à ceux qui sont passés : Soufian (pour les normes APAs), merci Rick (pour les galopins), merci Carla (pour les muffins), merci Kevin (de m’apprendre des trucs au quotidien), merci Lucie (d’être un rayon de soleil), merci Yuna (na ; pour ta tolérance face à mes idées fixes et d’être devenue une réelle amie), merci Noémie (pour tes précieux

conseils), merci Marine, merci Mohamed, merci Aurore, merci Florian et merci à toutes et tous plus généralement. Vous avez une place spéciale dans mon cœur.

Merci à mes amis, celles et ceux qui m'ont accompagné le long de cette route. Un merci aux psychopotes, et tout particulièrement à Lola. Un merci à Anna. Un merci à Camille, à Micka, et à celui qui nous rejoindra en cours de route. Un merci à Julien. Un merci à Cloé. À Inès. À Marie.

Un merci tout particulier à toute ma famille. Maman, je pèse mes mots lorsque je dis que rien de tout cela n'aurait été possible sans ton soutien. Solène, merci d'être une source d'inspiration quotidienne de combativité. Merci Sébastien d'être un exemple de gentillesse et de bienveillance. Merci Papa.

Un merci également à celles et ceux, en Ardèche, qui m'ont fait une place. Un nombre non négligeable des mots constituant ce manuscrit y ont été écrits. Merci Sandrine, merci Jean-Mi, merci Yoann. En voyant l'impact qu'ont eu ces dernières années sur le monde, je mesure la chance que j'ai eu d'avoir ce refuge.

Léa, merci pour tout. La thèse, c'est un long chemin, et je suis extrêmement chanceux de ne pas l'avoir fait tout seul. Je pense qu'avec le soutien que tu m'as apporté, tu mérites un peu aussi un titre de docteur. Maintenant, j'ai hâte de pouvoir décider quels seront les prochains qu'on décidera d'arpenter. Aussi, je suis content de savoir que le 一二三四五六七 ne sera jamais vraiment fini. Je te l'ai souvent répété, mais je doute de plus en plus que ces mots arrivent un jour à retranscrire à quel point je le pense, merci pour tout.

Résumé

Dans le champ de la cognition sociale implicite, les réponses évaluatives indirectes représentent une opportunité de dépasser certaines limites de l'auto-report. Sensées capter quelque chose de progressivement encodé au cours du temps et guidant nos comportements, ces mesures permettraient de déterminer l'attitude que quelqu'un a envers quelque chose, et ce, même lorsque cet individu ne voudrait pas ou ne pourrait pas dévoiler ses préférences. Pour la plupart des modèles théoriques rendant compte de ces réponses comportementales, c'est au travers d'expériences répétées que nous développons ces réponses évaluatives indirectes. De récents travaux expérimentaux, cependant, ont mis en évidence l'impact que de simples instructions pouvaient avoir sur ces réponses évaluatives. À travers ce travail de thèse, nous défendons l'idée que les effets de l'exposition répétée et ceux de simples instructions diffèrent. Pour investiguer cette question, nous avons tout d'abord développé un paradigme d'entraînement à l'approche et à l'évitement au cours duquel nos participant·es devaient approcher et éviter des stimuli de manière répétée. Après avoir montré que de nouvelles réponses évaluatives indirectes naissent de ce type d'entraînement (Exp. 1a–2), nous avons comparé ce paradigme expérimental à une procédure basée sur les instructions (Exp. 3–7). Parmi ces cinq études comparant les deux procédures sur plusieurs types de réponses évaluatives indirectes, auprès de différentes populations, et dans différentes situations, deux ont révélé une plus grande efficacité des entraînements à l'approche et à l'évitement (les trois autres ne s'étant pas révélées conclusives). Deux études complémentaires ont traité la question des théories naïves que pouvaient avoir des individus sur cette question (Exp. 8 & 9). Ces résultats, pris dans leur ensemble, s'inscrivent dans la continuité d'avancées théoriques récentes dans le champ de la cognition sociale implicite et nous amènent à recommander de privilégier des paradigmes tels que les entraînements à l'approche et à l'évitement comparativement aux paradigmes basés sur de simples instructions.

Mots-clés : Approche/évitement, Attitudes, Cognition sociale implicite, Évaluation, Mesures indirectes

Abstract

In the field of implicit social cognition, indirect evaluative responses represent an opportunity to overcome some of the limitations of self-report. Theoretically capturing something progressively encoded over time and guiding our behaviors, these measures would allow us to determine the attitude that people have towards something, even when these people would not or could not reveal their preferences. For most theoretical models accounting for these behavioral responses, it is through repeated experience that we develop indirect evaluative responses. Recent experimental work, however, highlighted the impact that simple instructions can have on these evaluative responses. Throughout this dissertation, we argue that the effects of repeated experience and simple instructions differ. To investigate this question, we first developed an approach-avoidance training paradigm in which our participants were asked to repeatedly approach and avoid stimuli. After showing that new indirect evaluative responses emerged from this type of training (Exp. 1a–2), we compared this experimental paradigm to an instruction-based procedure (Exp. 3–7). Of these five studies comparing the two procedures on several types of indirect evaluative responses, across different populations, and in different situations, two revealed greater effectiveness of approach and avoidance training (the other three were inconclusive). Two additional experiments addressed the issue of naive theories that individuals might have about this issue (Exp. 8 & 9). Taken together, these results are consistent with recent theoretical advances in the field of implicit social cognition and lead us to recommend paradigms such as approach and avoidance training over paradigms based on simple instructions.

Keywords: Approach-avoidance, Attitudes, Evaluation, Indirect measures, Implicit social cognition

Forewords

The research compiled in the present manuscript results from the collaboration between Cédric Batailler, Dominique Muller, and Cécile Nurra. This PhD work was funded by Université Grenoble Alpes with an Alpes Grenoble Innovation Recherche AGIR grant. The PhD was conducted at the LIP/PC2S laboratory. Some studies were conducted at Bertram Gawronski's Social Cognition Lab, University of Texas at Austin.

The present manuscript tackles the question of how performing approach and avoidance movements toward a target impact subsequent behaviors toward the said target. It specifically addresses the role of repeated experience in the emergence of the new behaviors. This theme is partly based on previous work initiated by Dominique Muller, François Ric, and Théodore Alexopoulos and pursued during Marine Rougier's PhD which resulted in the Visual Approach and Avoidance by the Self Task (the VAAST; Rougier et al., 2018)—a behavioral measure designed to robustly capture approach and avoidance tendencies. In the present manuscript, we adapted this task originally designed to measure approach and avoidance tendencies to a different context: approach and avoidance training. We used this novel task to assess the difference between repeated-experience- and instructions-based procedures when it comes to the acquisition of indirect evaluative responses. The present manuscript was written following on the APA7 recommendation.

Table of Content

	Page
<i>Acknowledgements</i>	<i>i</i>
<i>Résumé</i>	<i>vii</i>
<i>Abstract</i>	<i>ix</i>
<i>Forewords</i>	<i>xi</i>
<i>Table of Content</i>	<i>xiii</i>
<i>List of Tables and Figures</i>	<i>xvi</i>
 Preamble	 3
 PART I: THEORETICAL BACKGROUND	 9
 Chapter 1 • Attitudes: History, Usage, and Challenges	 11
The Early Works on Attitude	12
A Common Language for Attitudinal Research	13
On the Importance of Keeping Explanandum and Explanans Independent.....	15
Studying Evaluative Responses and Evaluation to Understand Behavior	16
The Functional-Cognitive Framework of Attitude in the Present Manuscript.....	21
Some Limitations of Self-Report Measures of Attitudes.....	22
Attitudes? What Do People Even Know about Them?	24
Attitudes? Would People Even Want to Express Them?	25
Beyond the Limitations of Self-Report	26
 Chapter 2 • The Revolution of Indirect Measures	 27
The Evaluative Priming Task.....	28
The Implicit Association Test	32
The Theoretical Origins of Implicit Social Cognition	37
The First Root of Implicit Social Cognition: Automaticity	38
The Second Root of Implicit Social Cognition: Awareness.....	40
The Current Landscape of Implicit Social Cognition	41
 Chapter 3 • Theoretical Models of Evaluation	 45
The Two Types of Processes in Dual Models.....	46
Associative Processes	46
Propositional Processes	50

Theoretical Models of Evaluation	53
The SEM.....	55
The APE Model.....	58
Single Process Propositional Models	63
Theoretical Models of Evaluation and the Role of Actual Experience.....	67
Chapter 4 • Observing Evaluative Learning	71
Paradigms Based on Repeated Experience	71
Evaluative Conditioning	72
Approach and Avoidance Training	76
Approach-Avoidance Training and Theoretical Models of Evaluation	80
Learning Through Instructions.....	83
Chapter 5 • Overview of the Present Dissertation	89
PART II: EXPERIMENTAL PART	93
Chapter 6 • A VAAST-based approach-avoidance training	95
Implicit Identification and the Gender Gap in Science	96
Increasing Math Implicit Identification with an Approach Training	98
Assessing the Evidential Value in Favor of the Math Approach Training	98
Increasing the Evidence for the Math Approach Training.....	100
Overview of the Experiments.....	103
Experiment 1a	104
Method.....	104
Analysis and discussion.....	109
Experiment 1b	112
Method.....	107
Results and Discussion	109
Experiment 2	118
Method.....	119
Analysis	122
Complementary analyses	124
General Discussion	125
Conclusion	129
Chapter 7 • Comparing Approach-Avoidance Procedures	131
Repeated Experience and the Acquisition of Indirect Evaluative Response	131
Approach-Avoidance Procedures.....	133
Capturing Evaluation with the Affect Misattribution Procedure	135
Experiment 3	136
Method.....	137
Results and Discussion	142

Experiment 4	146
Method.....	146
Results	149
General Discussion	152
Conclusion	154
Chapter 8 • Approach-Avoidance Effects in Incidental Designs	155
Approach-Avoidance Training Design	155
Incidental Approach-Avoidance Training.....	156
Improving Woud et al.'s Paradigm	158
Experiment 5	159
Method.....	161
Results	167
General Discussion	170
Conclusion	174
Chapter 9 • Effects of Approach-Avoidance Procedures Over Time	175
On the Stability of Indirect Evaluative Responses.....	176
Long-Term Encoding and Indirect Evaluative Responses.....	177
Approach-Avoidance Training to Sustainably Acquire Evaluative Responses.....	178
Experiment 6	180
Method.....	181
Results and Discussion	187
Experiment 7	196
Method.....	196
Results and Discussion	199
General Discussion	201
Conclusion	205
Chapter 10 • Lay Beliefs About Approach-Avoidance Procedures	207
The Effects of Approach-Avoidance Training and Approach-Avoidance Instructions....	207
On the Importance to Investigate People's Lay Belief	208
People's Lay Theories Regarding Approach-Avoidance Procedures	209
Experiment 8	210
Method.....	211
Results and Discussion	216
Experiment 9	218
Method.....	219
Results and Discussion	221
General Discussion	226
Conclusion	228

PART III: GENERAL DISCUSSION	229
Chapter 11 • General Discussion	231
Summary of the Results	231
Implications of our Results	235
Models of Evaluation	235
On the Effect Size of the Difference Between Approach-Avoidance Procedure.....	236
Other Research Focusing on the Contribution of Repeated Experience to	243
Evaluative Learning	
Methodological Limitations and Ways Forwards	257
Concluding thoughts	249
References	261
APPENDICES	297

List of Tables and Figures

Tables

Table 5.1	Overview of the Experiments	92
Table 7.1	Summary Statistics of Experiment 3's Mixed Model.....	145
Table 7.2	Summary Statistics of Experiment 4's Mixed Model.....	151
Table 8.1	Characteristics of the Stimuli Used in Experiment 5.....	160
Table 8.2	Example of Stimuli Used in Experiment 5	166
Table 8.3	Summary Statistics of Experiment 5's Mixed Model.....	170
Table 9.1	Summary Statistics of Experiment 6's Model	192

Figures

Figure 2.1	Time Course of the Evaluative Priming Task Trials	30
Figure 2.2	Category Configurations in the Different Blocks of an IAT	34
Figure 2.3	Block Structure in an IAT.....	35
Figure 3.1	Information in an Associative Network.....	48
Figure 3.2	Representations of Associative and Propositional Relationships	51
Figure 3.3	Importance of Associative Processes in Indirect Evaluation.....	55
Figure 3.4	Processes in the SEM and the APE model	60
Figure 3.5	A Single-Process Propositional Model	65
Figure 4.1	Example of Advertisement	74
Figure 4.2	Trials in a Black Face Approach Training.....	79
Figure 6.1	Time Course of a VAAST Trial in Experiment 1a.....	108
Figure 6.2	Results of Experiment 1a.....	111
Figure 6.3	Results of Experiment 1b.....	117
Figure 6.4	Results of Experiment 2.....	123
Figure 6.5	Meta-analysis of the Effect of Math Approach Training.....	125
Figure 7.1	Stimuli Used in Experiment 3.....	139
Figure 7.2	Time Course of Experiment 3 According to the Experimental Condition	139
Figure 7.3	Time Course of an AMP Trial in Experiment 3	142
Figure 7.4	Results of Experiment 3.....	144

Figure 7.5	Example of Stimuli Used in Experiment 4	148
Figure 7.6	Results of Experiment 4.....	151
Figure 8.1	Time Course of Experiment 4 According to the Experimental Condition	162
Figure 8.2	Time Course of an AMP Trial in Experiment 5	165
Figure 8.3	Results of Experiment 5.....	169
Figure 8.4	Relationship Between the Complexity of a Proposition and Probability	173
	of Acquisition	
Figure 9.1	Single-Process Propositional Model of Evaluation	177
Figure 9.2	Time Course of Experiment 6.....	183
Figure 9.3	IAT Block Configuration in Experiment 6	186
Figure 9.4	Flowchart of Participants in Experiment 6	188
Figure 9.5	Duration Between Stage 1 and 2 in Experiment 6.....	189
Figure 9.6	Results of Experiment 6.....	191
Figure 9.7	Sensitivity of Experiment 6's Results to Sample Size.....	195
Figure 9.8	Sensitivity Analysis for Experiment 7	197
Figure 9.9	Flowchart of Participants in Experiment 7	199
Figure 9.10	Results of Experiment 7.....	201
Figure 10.1	Scaled Used in Experiment 9.....	215
Figure 10.2	Results of Experiment 8.....	217
Figure 10.3	Results of Experiment 9.....	223
Figure 10.4	Distribution of the Estimated Outcomes in Experiment 9	225
Figure 11.1	Reported Sample Size in Social Psychology Journals.....	241
Figure 11.2	Classification Images of Approached and Avoided Faces	247

On the contribution of repeated experience to evaluative learning: An approach-avoidance procedure investigation

Imagine starting a new job in an organization. You don't know anyone there, but you do know that you'll have to quickly find your feet if you don't want this experience to become unpleasant. To start, few friendly colleagues could make your days easier. Workplace can be a scary world and coffee breaks in nice company would sure help you through the day. The whole problem here is that you're an introvert. It costs you a lot to reach out. You know that you're the type of person that needs a few days on their own after 5 minutes of small talk, just so you can recover—the type of person that needs to repeat to themselves their own name to be sure that they won't forget it when asked. Finding the right people at this new place will be the main challenge of your first days there.

You arrived on Monday at the sad tower where your office is. On your way there, you went to a small hole in the wall selling sandwiches, just to avoid meeting people at the cafeteria. Someone was cheerfully chatting with the vendor. You chose the less depressing snack. You put back your earphones for the two-block walk separating the sandwich place from your office, took your badge at the reception, the elevator to the 26th floor where you met the HR department, and went straight to your office. On your way there, you bumped into the person that you are replacing, a person who looked absolutely ready to enjoy their well-deserved retirement. The time passed, you filled some excel sheets, ate your not-depressing sandwich, spent some time looking at the window, and, at around 5p.m., called it a day. On your way to the elevator, you heard some laughs by the copier, turned your head, and recognized the sandwich-place customer. You took the bus. It was your first day at this new job.

The next few days, in addition to doing some more spreadsheet filling, you learned that the sandwich-place person was named Alex. You saw Alex joking around with other employees, you saw Alex taking a tomato soup at the vending machine, and you heard Alex humming what appeared to your ears as a tribute to ABBA. You decided that this person was probably nice, and maybe, in a few weeks, you would have gathered enough resource to drop by and ask Alex to join you on your way to the not-depressing sandwich place. You designed this plan after a two-hour session of spreadsheet filling and left with a sense of accomplishment. But the very next day, when you arrived at the sad tower, a series of events disemboweled your carefully crafted strategy. You entered the lobby, glancing at your shoes, and, halfway through, lifted your chin up. In front of you, stepping into the elevator, was Alex. You now have to decide whether you will speed up to take this elevator. You will have to quickly decide whether you want to, as an introvert, take a 20-ish floor ride with an almost stranger who probably will recognize you and who will probably engage in conversation. Walking slowly would be easier. You would avoid a stressful situation. Plus, it would let you finish this podcast on ships from the 18th century. Yet, you can see yourself walking faster. You might have in front of you the opportunity to find a friend in the workplace. A small behavior.

Now imagine that, on your first day, you didn't stop at the sandwich place. You didn't hear anyone joking around, taking tomato soup, or humming ABBA. Rather, when leaving their office with their box full of supplies, the person that you're replacing told you: "Oh, and by the way, Alex's nice, I'm sure that you'll get along. Look for a small person, dark hair, glasses, and a red bike helmet". Would this information trigger the same kinds of question in your head? Would it trigger the same behaviors? Is this information enough for you to speed up, and decide that taking this elevator is the right choice?

The question that we will address in the present dissertation is the same that emerged when we wonder how our past experiences could shape our reactions toward Alex. Do repeated experiences add something over merely being told something? In the present dissertation, however, rather than elevator-related behaviors, we will focus on one type of behavior, called indirect evaluative responses. These indirect evaluative responses are the result of research effort in social psychology and are thought to be particularly diagnostic of spontaneous behaviors. We will ask whether paradigms based on repeated experience are more likely to make someone acquire a specific indirect evaluative response compared to their mere instructions-based counterparts. We will be focusing specifically on approach-avoidance training.

The first part of this dissertation will be theoretical. We will undertake a somewhat extensive review of the literature relevant to our research question. To offer a comprehensive view of why it is important to assess the contribution of repeated experience regarding the acquisition of indirect evaluative responses, we will introduce a core concept of the social psychology literature: the concept of attitude (**Chapter 1**). We will use this chapter as an opportunity to offer some clarity about the language that we will use. We will then dive into one of the subfields of social psychology, the subfield of the indirect evaluative responses that we mentioned above: the subfield of implicit social cognition (**Chapter 2**). We will first discuss how implicit social cognition emerged as well as the first measures developed in this context: the evaluative priming task and the IAT. After introducing these measures (which belong to the category of indirect evaluative response), we will see how ideas from cognitive psychology facilitated the advances observed in implicit social cognition. In **Chapter 3**, we will review how the implicit social cognition field evolved with the modern theories of evaluation acquisition. We will be particularly focusing on the theories which give a special

place to indirect evaluative response acquisition: the system of evaluation model (the SEM),¹ the associative-propositional evaluation model (the APE model), and the single-process propositional models. This section in particular will be helpful to understand the importance given to repeated experience in the learning of indirect evaluative responses. Finally, we will review some of the paradigms that have been used to explore the question of evaluation acquisition, including the paradigms based on repeated experiences and their counterparts which are not (**Chapter 4**). In these sections, we will highlight why we think that repeated experiences do contribute to the acquisition of indirect evaluative responses, and why we think that it is an important issue to address.

As we will see in more detail later (see **Chapter 5**), our experimental work will be organized according to the different objectives of our experiments. We will first develop an approach-avoidance training based on recent advances in the implementation of approach-avoidance movements (**Chapter 6**). Then, we will compare this approach-avoidance training to an approach-avoidance instruction procedure. We will compare the two in different situations and using different indirect evaluative responses (**Chapters 7–9**). Finally, in order to provide the best recommendations we can regarding the use of approach-avoidance procedures, we will focus on people's lay beliefs about the effectiveness of approach-avoidance training and instructions (**Chapter 10**).

Our final chapter (**Chapter 11**) will be dedicated to the general discussion. In this chapter, we will come back to the theories that we reviewed while laying the theoretical foundation for this dissertation. We will also discuss our results with regards to the lines of research that investigated the differences between repeated-experience- and instructions-based paradigms of implicit social cognition. Finally, we will consider ways to improve

¹ A table of abbreviations used in this dissertation can be found in Appendix 1.

further research on this topic. Ultimately, the present dissertation will represent the opportunity to address a question with both theoretical and applicative implications. This question is the one of the contributions of repeated experience in some of the paradigms used in implicit social cognition.

PART I: THEORETICAL BACKGROUND

Chapter 1

Attitudes: History, Usage, and Challenges

One of the key concepts adopted by social psychologists can be traced back to a chapter in *The Handbook of Social Psychology* (1935). In this chapter, Allport (1935) reviews the use of the term *attitude* and provides a definition in terms of “neuropsychic state of readiness for mental and physical activity”. This definition was the first step on a journey to better understand consistency in human behavior. As reported by Allport (1935), at the time, it was common for scientists interested in understanding behavior to debate on whether human behaviors were a result of heredity or environment. By introducing and putting forward a concept for which no origin was assumed, merely describing a probability to act a certain way, social psychologists could start to focus on the factors making people likely to act in a predictable manner, no matter whether these factors were internal or external (Kelley, 1973). Allport’s review on attitudes (1935) is often perceived as the seminal paper that helped popularize the use of attitudes among the social psychologists. However, at the time, this concept was facing many challenges.

An illustration of one of these challenges can be illustrated by a study conducted to investigate prejudice. In the ‘30s, LaPiere (1934) tried to capture attitudes—the state of readiness to act in a certain way described by Allport (1935)—to see whether it was possible to predict people’s behavior. Along with a Chinese couple, he spent two years traveling across the US, going from hotel to hotel. When their trip came to an end, LaPiere mailed one question to the hotels they visited: Would they accept “members of the Chinese race” in their establishment? The result of this field study could have marked the end of attitude research: While 92% of the establishments responded they would refuse to welcome Chinese

individuals, none but one establishment actually refused to welcome LaPiere and the accompanying couple. The problem illustrated here is actually simple to understand: It is hard to know how to capture the state of readiness people have. As later phrased in the literature, “Asking people to report on their attitudes, will almost always result in an answer—but it often remains unclear what exactly this answer means” (Schwarz, 2008, p. 49).

The Early Works on Attitude

Years passed and some researchers argued that attitudinal research in social psychology was a dead end (e.g., Wicker, 1969). It was only when researchers focused on building theories of attitudes that this field of social psychology gained in popularity. Researchers started theorizing how attitudes worked and what their relationships with behavior were. The reflections soon translated into principles which are well known to attitudinal researchers, such as the correspondence principle. If someone has to report their attitude toward something, the closer the attitudinal question is from behavior in terms of generality-specificity, the higher the correlation between the attitudinal question and the behaviors (i.e., the degree of correspondence between attitudinal and behavioral entities is important; Ajzen & Fishbein, 1977). For example, asking people their general views on the use of contraceptives (something general) should only be able to predict weakly their use of condoms (something specific). This example might seem trivial to the reader, but, as we will review, many not-so-trivial reflections came from the social psychology literature regarding how we should capture attitudes and when these attitudes are supposed to influence behaviors.

In the present dissertation, we will focus on some of the theories which were developed in the field of attitudinal research, as well as the instruments related to these theories. These theories and measures emerged from what is now called the implicit social cognition literature—a subfield of social psychology which emerged in the 90s (Greenwald &

Banaji, 1995; see also, Gawronski & Payne, 2010). Specifically, we will focus on some of the assumptions related to the role of repeated experience which were embedded in the theories of implicit social cognition. Later, we will discuss why this subfield emerged, how it emerged, what principles emerged from this literature, and what kind of expectation one can have based on those principles. Beforehand, however, we will take the time to formalize how we will discuss attitudinal research. Indeed, many definitions coexist in the literature (Bohner & Dickel, 2011), and we think that the more precise we are, the lower the risk for what we offer to become impossible to falsify.

A Common Language for Attitudinal Research

As we just saw, attitude is one of the concepts used by social psychologists to understand consistency in the behaviors of human beings. Initially defined as “a mental and neural state of readiness, organized through experience, exerting a directive or dynamic influence upon the individual's response to all objects and situations with which it is related” (Allport, 1935), attitudes would help understand and predict why we act one way or another. Decades after its initial introduction, the attitude concept now often translates as a mental construct which has a significant causal impact on one’s behavior. Attitudinal research is oriented to understanding how attitudes are formed and when they will be activated (i.e., how attitudes appear, when they will influence behavior). It has to be noted that the advances in attitudinal research resulted in many theoretical models, each endorsing a different definition of attitude (see Bohner & Dickel, 2011). This situation follows Nolan Bushnell’s Law: While it is easy to intuitively see what an attitude refers to (“Do you like ice cream or not?”), there is a lot of debate on what a good definition of attitude is.

In the present dissertation, focusing on the contribution of repeated experience will lead us to explore the predictions of several theoretical models of attitude. Because each of them adopts different assumptions regarding what an attitude is, we will refrain from making

this term the core concept of this dissertation. Instead, we will use a meta-theoretical framework allowing us to describe the many research conducted by attitudinal researchers in a common vocabulary. In doing so, we hope to avoid a situation where we would endorse theoretical assumptions without being explicit about it. This meta-theoretical framework, initially described by De Houwer et al. (2013), was designed with the goal to provide a language for researchers to talk to each other, as well as to make it easier to avoid common pitfalls of attitudinal research. In the present section, we will describe the concerns that De Houwer et al.'s (2013) meta-theoretical framework is addressing, and the perspective it adopts to do so.

Before describing one of the issues addressed by De Houwer et al.'s (2013) meta-theoretical framework, we will first have to define some vocabulary. Two terms will be especially useful: *explanans* and *explanandum*. This vocabulary, related to the idea of explanation, targets two different things. The concept of *explanandum* describes something that researchers are interested in understanding—in explaining. Storms (or human behaviors), for example, could be something researchers want to understand; it could be an *explanandum*. To understand something, researchers usually describe processes whose goal is to explain the phenomenon. In the context of weather, to pursue with this example, scientists could explain storms as a result of a process involving differences in atmospheric pressures (or human behaviors as driven by attitudes). The processes explaining our object of interest are called the *explanans*. In attitudinal research, the thing that researchers are interested in predicting—the *explanandum*—is the human behavior and, to better understand the human behavior, the recruited concepts—the *explanans*—are attitudes. A useful mnemonic is that *explanans* sounds more like “explaining” than *explanandum*, and the *explanans* does refer to the things *explaining* a phenomenon.

On the Importance of Keeping *Explanandum* and *Explanans* Independent

For a theoretical model to be efficient, it is especially important to keep the *explanans* and *explanandum* distinct from each other. One of the pitfalls in attitudinal research is that the *explanans* and *explanandum* are often conflated with each other. In other words, for De Houwer et al. (2013), it is sometimes hard to know whether something is a behavior that researchers are interested in understanding (the *explanandum*) or an attitude (the *explanans*). To understand this idea, we will focus on a practice which is common among attitudinal researchers: using behaviors as a proxy for attitudes. For example, a researcher interested in math-related behavior could use the behavioral response of a participant on an evaluative rating scale (“from one to ten, how much do you like math?”) as a proxy for attitudes toward math.² Some assumptions, however, need to be made so that this specific behavior can be considered as a good proxy.

Here, in the math example, researchers are inferring the existence of an attitude (X) from a behavior (Y). Yet, the core assumption of attitudinal research is that attitude (X) influences the behavior (Y). In formal logic, inferring the existence of X from Y based on the premise “if X, then Y” is known as affirming the consequent. It is considered as a fallacy as soon as Y varies because of factors other than X. In attitudinal research, there is plenty of evidence showing that an attitudinal measure can vary because of factors other than a change of attitude (e.g., social desirability; DeMaio, 1984). In our example, to draw any conclusion regarding the link between attitudes toward math and math-related behavior, we would have

² Note that this theoretical framework endorses some of the assumptions of radical behaviorism. Notably, anything that can be observed is a behavior (even once own thought). Both answering a questionnaire and decided to follow a math curriculum are behaviors (Skinner, 1953).

to assume that there is no factor other than attitude which influences the attitude measure. It is easy to imagine that the answer to the measure assessing how much people like math depends on who asks the question. In a context where the mental mechanisms used to explain a phenomenon are the one that needs to be explained, attitudinal measures become a behavior just like any other (De Houwer et al., 2013). Many researchers have outlined this recurring problem in attitudinal research (e.g., Fazio, 2007; Krosnick et al., 2005).

In their meta-theoretical model for attitudinal research, De Houwer et al. (2013) offer a vocabulary which makes it easier to keep *explanandum* and *explanans* independent from each other. In adopting this vocabulary, besides avoiding some of the pitfalls of attitudinal research, we can describe any of the theoretical models of attitude in common terms. In their meta-theoretical framework, De Houwer et al. (2013) suggest that researchers should focus on evaluative responses (the *explanandum*) and models explaining evaluation (the *explanans*). In the next section, we will introduce these two notions which will be used in lieu of the traditional attitude concept.

Studying Evaluative Responses and Evaluation to Understand Behavior

Evaluative Responses, or the Functional Level

In De Houwer et al.'s (2013) meta-theoretical framework, the *explanandum*—the phenomena that researchers are interested in explaining—are evaluative responses. Evaluative responses are a behavior that can be defined according to two principles: scientific conventions and correlations with other evaluative responses. With the first principle, if it is consensual among scientists that circling a number on a questionnaire asking “How much do you like math?” is an evaluative response, then it will be considered as an evaluative response. With the second principle, if a behavior correlates with one which belongs to the category of the evaluative responses, then it will also be considered as an evaluative response. At first glance, defining the *explanandum* in such a manner appears problematic because it

seems arbitrary. However, with regard to the conflation between the *explanandum* and the *explanans*, this definition is actually useful.

As we discussed above, the *explanandum* and the *explanans* are conflated when researchers equate behavioral observations with mental constructs, as it is the case when someone considers that an answer to the “Do you like math?” is an attitude (De Houwer et al., 2013). Here, researchers would use a behavior as a proxy for attitudes; the same attitudes that are supposed to account for behaviors. Such recursive logic is what we are trying to avoid as it creates dilemmas impossible to solve. Note that these considerations generalize to any situation where a behavior would be used as a proxy for a mental concept aimed at explaining a phenomenon (e.g., spreading-of-alternative as an index for cognitive dissonance; Brehm, 1956; De Houwer et al., 2013). Focusing on evaluative responses does not pose that kind of problem because no behavior is defined by mental mechanisms that underlie them.

As we just said, in De Houwer et al.’s (2013) meta-theoretical framework, the *explanandum* do not embed any cognitive assumptions in their definition. Yet, it is possible to gain knowledge related to human behavior. This knowledge will fit in what De Houwer et al.’s (2013) call a *functional level of explanation*. In this functional level of explanation, it is possible to build knowledge by investigating causation in evaluative responses. Indeed, there is no need to assume mental mechanisms to be able to observe, for example, that a change in the environment results in a change in the evaluative response (e.g., Wittenbrink et al., 2001) or that an evaluative response predicts another (e.g., Necka et al., 2015). Mere experimental method is usually enough to establish causation in psychology (but see also Rohrer, 2018). Here, the functional level of explanation is called so because it allows researchers to describe evaluative responses *as a function* of a factor. Note that the functional level has nothing to do with something functional from an evolutionary perspective. By comparing evaluative responses that an individual produces in the presence of a particular stimulus compared to a

situation where the stimulus is absent, researchers can establish that the said stimulus influences have a causal impact on the evaluative response. This ability to find causal effects in the environment is critical to foster research related to evaluative responses.

Still, one can have the feeling that relying on this functional level of explanation would be equivalent as drawing arrows between stimuli and behaviors—that is, behaviorism. One could argue that the time has passed when we considered the human mind as an inaccessible black box. This view notably accompanied the development of cognitive psychology (see Miller, 2003): a lot of the progress observed in modern psychology came from the modeling of unobservable cognitive mechanisms. So how does the meta-theoretical framework offered by De Houwer et al. (2013) overcome the limitations of behaviorism?

Evaluation, or the Cognitive Level

At first glance, the described meta-framework promotes a black-box-like vision of human mind in which researchers can do nothing but observe relationships between stimuli and responses as the only means to understand behavior. This argument, however, only holds if we ignore the second level of explanation this framework offers. In the meta-theoretical framework of evaluation (De Houwer et al., 2013), researchers are interested in both a functional and a cognitive level. The second level, the cognitive one, describes where the cognitive processes underlying evaluation will be formalized—evaluation referring to the effect that factors have on evaluative responses. The same way functional relationships between evaluative responses were the main focus of the functional level, cognitive explanations of evaluation will be the main focus of the cognitive.

Researchers consider an evaluation to have happened when they can demonstrate that an evaluative response is caused by a specific factor (for example, when they can demonstrate that the response to the question, "How much do you like math?" is affected by who asks the question). Critically, modern cognitive psychology argues that every impact a

stimulus has on behavior must be the result of mental representations produced and reactivated via certain mental processes (contrary to an explanation which would rely on something happening outside of one's cognitive system). At the cognitive level, researchers will formalize the processes mediating the effects of various factors on evaluative responses (i.e., evaluations; De Houwer et al., 2013). In such a framework, attitudinal literature aiming at explaining the circumstances under which attitudes influenced behavior as a collection of hypotheses regarding evaluation. These hypotheses address questions regarding why a stimulus can cause a specific evaluative response or why this evaluative response can change depending on factors in the environment.

Note that while the cognitive level of explanation is purely theoretical and will never be directly observable, it is still possible to assess its validity (Gawronski & Bodenhausen, 2015). In order to be useful, a cognitive theory must organize the phenomena observed at the functional level of explanation accurately and must allow researchers to make new predictions regarding evaluation. Any theory failing any of these two objectives would eventually disappear. Together, the functional level and the cognitive level form a meta-theoretical model of evaluation.

A Framework to Study Attitude

These two levels of explanation, when combined, form a meta-theoretical framework that researchers may employ to study behavior without conflating the *explanandum* and the *explanans*. Critically for the present dissertation, this meta-theoretical framework's scope will allow us to describe several models in a common vocabulary. These models were initially framed as models of attitude, however, any research subject which has been the focus of attitudinal research can be reframed in this meta-theoretical framework (De Houwer et al., 2013). Research on acquisition of new attitudes can be reframed as the investigation of the effect of stimulus history on evaluative responses, research on activation of attitudes can be reframed as the investigation of the effect of stimulus context on evaluative responses, and research on the effect of attitudes on behavior can be reframed as the investigation of the varying nature of evaluative responses toward a stimulus (De Houwer et al., 2013). Contrary to traditional attitude research, however, this meta-theoretical framework is designed to keep the things it has to explain and the concepts used to account for them—the *explanandum* and the *explanans*—clearly separated.

While distinct, the functional and cognitive levels of explanations mutually support each other. In the words of De Houwer et al. (2013), the knowledge gained at the functional level (i.e., knowledge about how evaluative responses relates with each other and what affects them) represents the foundation of the cognitive level—the data that the cognitive level of explanation will have to account for. A better understanding of evaluative responses at the functional level therefore represents a higher level of constraints on the theories formalized at the cognitive level, which usually results in more precise hypotheses. Any progress at the functional level translates in more precise theories at the cognitive level and any progress at the cognitive level of explanation translates in theories more effective at

organizing what we know about the factors driving evaluative responses (De Houwer et al., 2013).

At this stage, readers could think that this meta-theoretical framework looks a lot like the distinction between theory and operationalization adopted by cognitive psychology (see Braisby et al., 1996). We generally agree with this observation, but think that taking the time to formalize this distinction in a meta-theoretical framework will help us organize the literature (De Houwer et al., 2013). A recent example showing that this message still benefits from being disseminated comes from the implicit bias literature (which we will cover in more details later in the dissertation). Implicit bias is sometimes used to refer to an unconscious and invisible force driving people's behavior in an undesirable manner (e.g., an invisible force which would make us picture a male professor when someone is merely referring to a professor; Amodio & Mendoza, 2010). De Houwer (2019) recently suggested moving away from this definition to adopt one where implicit bias would be defined by the behavior rather than by hypothetical mechanisms underlying them. The proposed definition has notable implication regarding the public policy aiming at reducing implicit bias. As discussed by De Houwer (2019), there is a major advantage in trying to implement policies aiming at reducing problematic behaviors rather than targeting mechanisms about which researchers are not even certain.

The Functional-Cognitive Framework of Attitude in the Present Manuscript

Altogether, this meta-theoretical framework forces the theorists to be strict regarding what they are talking about. Albeit lengthy, we consider that the current section is essential before reviewing the attitudinal research to avoid confusion which exists in the literature and which recently forced researchers to ask for more clarity (e.g., Corneille & Hütter, 2020). By adopting this meta-theoretical framework, we limit the risk of relying on theoretical assumptions without being explicit about it. We will adopt De Houwer et al.'s (2013)

framework, therefore keeping the functional and cognitive level clearly separated. We will limit theoretical developments to the cognitive level of explanation.

Now that we described the meta-framework and explained why we decided to adopt it, we will come back to some of the challenges that attitudinal researchers were facing. This is when attitudinal researchers decided to use the advances of cognitive psychology to overcome some limitations of traditionally used attitudinal measure that the field of implicit social cognition is born. We will first discuss the challenges, before moving to a section dedicated to the early days of implicit social cognition. Critically, it is in the very root of this subfield that we will find some assumptions regarding the role of repeated experience.

Some Limitations of Self-Report Measures of Attitudes

When the first work which would later result in the birth of implicit social cognition appears, during the late '80s, one of the most popular tools among attitudinal researchers was self-report. If we use the meta-theoretical framework of evaluation that we adopt in the dissertation to describe the situation at the time, on a cognitive level, it was assumed that most people could hold inclinations, feeling, preconceived notions, or convictions toward something. Attitudes were the total sum of these cognitive contents (Thurstone, 1928). Research on attitudes allowed researchers to investigate societal issues such as intergroup relationship (Brewer & Kramer, 1985) and theories were developed to see how methods like persuasion could shape our attitudes (Petty & Cacioppo, 1986). For researchers, capturing attitudes was a useful manner to study behavior, as attitudes could sometimes influence behavior.

As we said, the most typical approach to capture the attitude someone could hold toward something was simply asking this person a direct question (we will use the term direct evaluative response to describe these questions relying on participants adopting a self-report perspective). Researchers interested in voting behavior in the US, for example, could ask

their participants how much they like Ronald Regan (or, for an up-to-date version, Donald Trump). For this question to be considered a valid attitudinal measure, however, we must make a few assumptions regarding what needs to happen in the respondent mind. When someone is facing this question, ideally, they would first make sense out of the question. They would then retrieve the relevant information in memory, consider what would be the most appropriate response, and, finally, report it (Schwarz, 2008). If participants cannot make sense of the question, if they cannot retrieve the relevant information in memory, if they are unable to accurately map their response onto the questionnaire, or if they cannot report their answer, the direct question asked should fail to capture an attitude. Assuming that every step just described occurs exactly as specified, researchers could use participants' answer and use it as a proxy for attitudes (for a brief discussion on different approaches of measurement, see John & Benet-Martinez, 2000).

But, as reviewed by Payne and Gawronski (2010), concerned about these processes emerged in the literature. Specifically, researchers were wondering whether people were really able to retrieve relevant information in memory and are whether they would always be willing to report their answers to the questions. These concerns, and how psychologists handle them, resulted in the implicit social cognition. We will describe in the following sections how these concerns were expressed at the time, starting with people's lack of introspective access and continuing with the control people can exert on their answers.

Attitudes? What Do People Even Know about Them?

Remember, one of the main assumptions of modern cognitive psychology is that every impact a stimulus has on behavior is the result of mental representations produced and reactivated via internal mental processes. In one of the most important work of social psychology, Nisbett and Wilson (1977) highlighted a limit related to people's ability to access internal mental processes, a limit which would apply to evaluative responses based on self-reports. The setup for two of the studies reported in this paper was pretty simple. In a commercial establishment, participants were presented a consumer survey. They were invited to rate several articles of clothing (Study 1: four different nightgowns; Study 2: four pairs of nylon stockings) by picking their favorite one. Once done, they were then invited to explain why they picked this article. Out of the 378 participants who took part in the studies, none mentioned the position of the article on the table as a reason for preference. Yet, results revealed that the articles on the right side were preferred compared to the one on the left side. An even more interesting observation from these studies is that, when participants were explicitly asked whether the location of the article on the table could have played a role on their preference, Nisbett and Wilson (1977) report participants unanimously denied it.

Since these two studies were conducted, they served as the textbook explanation of an important idea: People often lack the introspective access to the high-order processes which govern their mind. And this lack of introspective access is especially concerning when it comes to asking people what their attitudes toward something are. Could it be that the processes governing people's behaviors are not accessible by introspection? Could it be that what people report simply do not matter? This lack of introspective access could explain the low predictive validity of attitudes measure in some cases (e.g., Epley & Dunning, 2000) and has been considered as one of the limits of self-report measures.

Attitudes? Would People Even Want to Express Them?

Still, even if we consider that people can accurately identify the processes guiding their behaviors, another question exists: Would they want to report it? One could simply decide not to disclose what their attitude toward something is. Let's assume a situation where you hold very strong (and unjustified) negative attitudes toward ratatouille, a French stewed vegetable dish. It is likely that, if you met your in-laws for the first time, and they decided to go with ratatouille, your answer will be "Of course it does, I love ratatouille", when asked if it suits you. When someone has to disclose their attitude, different goals can compete in the answer-generation process (for a discussion, see E. E. Jones & Sigall, 1971). Among others, the motivations involved when answering a question can include social desirability—that is a tendency to give a favorable picture of oneself (for a review, see DeMaio, 1984). This kind of motivation can arise because of phenomena such as evaluation apprehension (the fear that your in-laws will judge you for your dislike of ratatouille; Rosenberg, 2009) or experimenter demand (your desire to go in the direction of your in-law's expectations; Orne, 1962). In terms of answer generation process, these kinds of motivations result in a change in participants' initial response. As soon as the respondent's motivation is different from "accurately report what I think the correct answer is", the measure designed to measure attitude should decrease in validity.

For this reason, social psychologists tried to develop paradigms to increase participants' motivation to accurately report what they think the answer should be. A notable example was the bogus pipeline (E. E. Jones & Sigall, 1971). In this paradigm, participants were convinced that a physiological device was able to predict the direction and amplitude of

the participants' emotional response³. Given that participants thought that the researchers would know if they hid something, the bogus pipeline paradigm made participants more likely to express negative feelings. This paradigm was soon followed by others.

Beyond the Limitations of Self-Report

At the same time concerns emerged regarding the use of so-called self-measure of attitudes, researchers developed new kinds of experimental paradigms. These paradigms, among which we can find the evaluative priming Task or Implicit Association Test (Fazio et al., 1986; Gawronski et al., 2020; Greenwald et al., 1998), represent the foundations of what we called earlier implicit social cognition. Because these new paradigms had the potential to overcome some of the aforementioned limits of self-report, they became extremely popular. The relevance of these events in the present dissertation lies in the fact that, alongside the instruments, came new theories. And, as we will see, these theories granted a special place to the principle of repeated experience. In the next chapter, we will first introduce the two paradigms who contributed the most to the early days of implicit social cognition and then discussed the cognitive psychology literature which inspired the researchers who developed these paradigms. We think that this chapter will later help the reader understand some of the assumptions of the modern theories of implicit social cognition as well as the kinds of paradigm used to investigate how people acquire patterns of behavior on evaluative priming task or IATs.

³ Note here that this paradigm led participant to focus on the emotional component of their response, therefore assuming that this emotional component is an important determinant of what their answer should have been without any motivation other than accuracy. This assumption is a strong one and is discussed in E. E. Jones and Sigall's (1971) original paper.

Chapter 2

The Revolution of Indirect Measures

As reviewed in the previous chapter, some researchers decided to focus on attitudes to better understand human behavior and to do so, they developed measures of attitude. In the present dissertation, we refer to these measures as evaluative responses (De Houwer et al., 2013). During the late '80s, some of these researchers interested in attitude developed new kinds of evaluative responses. These new measures—often referred to as *implicit measures* for reasons we will discuss later (see Gawronski & Hahn, 2019)—could potentially overcome some of the limits of self-report measures. The measures inspired by advances in cognitive psychology will represent the main focus of this manuscript. Note that we will not use the term *implicit measures* to describe this new category of paradigm. Rather, we will adopt the term indirect evaluative response.

The adoption of indirect evaluative measures produced insightful data on evaluation, suggesting that the nature of indirect evaluative responses differs from the nature of the direct evaluative responses. As reviewed by Gawronski et al. (2020), the dissociation between direct and indirect evaluative measures was shown with different kinds of empirical data. Notably, the correlation between indirect evaluative responses and direct evaluative responses tends to be rather low (see Cunningham et al., 2001; Hofmann et al., 2005), direct and indirect evaluative responses predict different kinds of behaviors (see Greenwald et al., 2009), and the same factor can impact direct and indirect evaluative responses differently (see Forscher et al., 2019). Altogether, these observations helped popularize indirect evaluative measures. At first glance, researchers successfully opened a new window into the human

mind. In the present dissertation, we are focusing on the origins of the behaviors observed through that window.

In the following sections, we will review two measures who belong to the category of the indirect evaluative responses and whose development resulted in the birth of implicit social cognition. After reviewing these two tasks, we will briefly review the literature that inspired them. While there are many different indirect evaluative responses (for a review, see Gawronski & Hahn, 2019), we decided to focus on the evaluative priming task and the IAT. These two measures were the first to be developed, and few others inspired the theories of implicit social cognition like they did. We consider this part important, because assumptions from this original literature were embedded in the implicit social cognition literature and can now be retrieved in modern theories. We will first review the evaluative priming task and then the IAT, keeping the functional level of explanation and the cognitive one separated from each other.

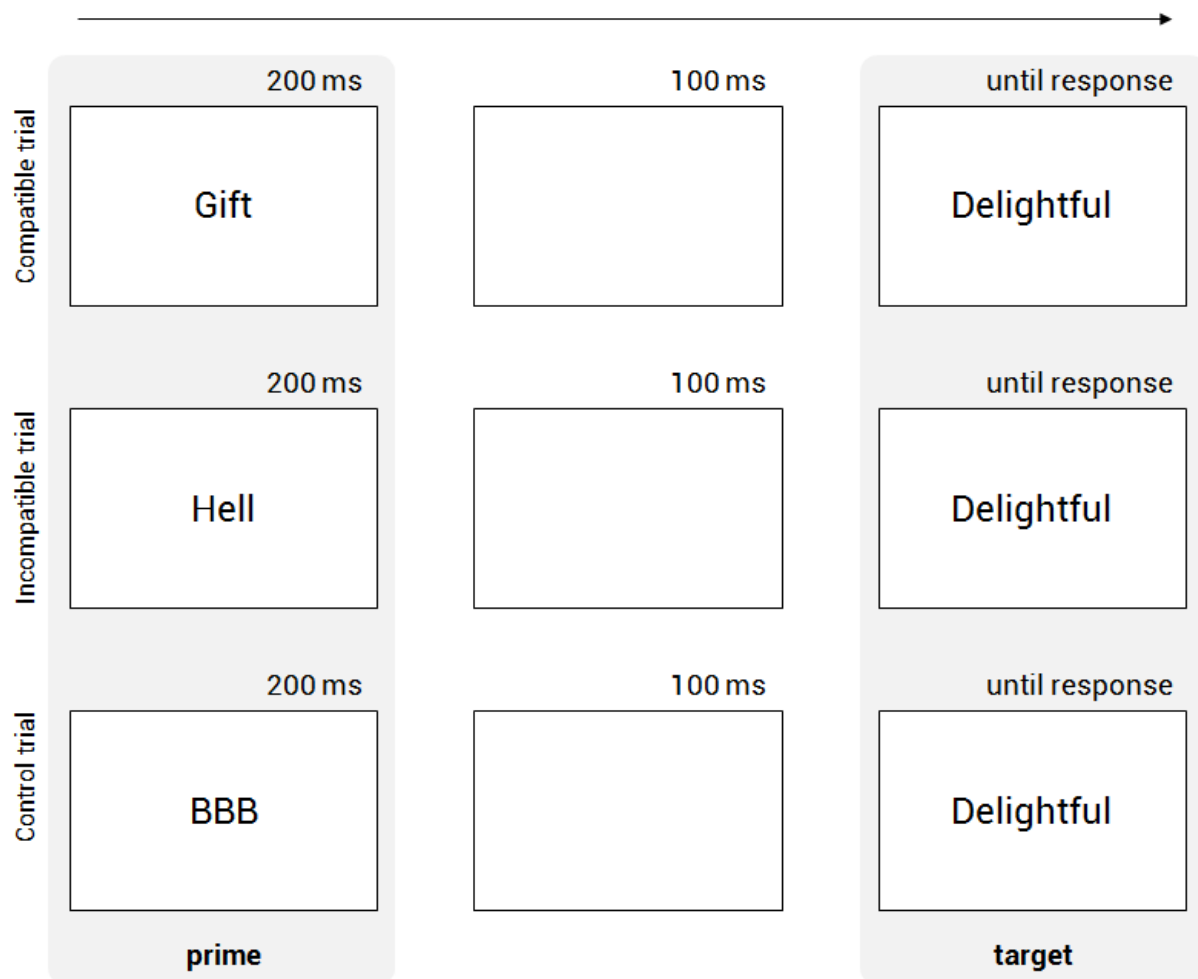
The Evaluative Priming Task

Historically, the evaluative priming task appeared before the IAT (Fazio et al., 1986). To introduce the reader to this task, we will first discuss what it is assumed to capture and, to do so, we will have to briefly discuss some of the theoretical assumptions making the evaluative priming task meaningful. Afterwards, we will describe what a participant completing an evaluative priming task does and detail some examples of application of this task.

First, to understand this task, we must discuss the assumptions made at the cognitive level. The evaluative priming task is best explained if one assumes the existence of *concepts* in the human mind. Concepts are merely a mental representation of an object, an idea, or even a social group and they are composed of different components, including an affective one (which can be either positive or negative). The *cat* concept, for example, can have a

positive affective component. When a concept possesses an affective component, it is likely that the positive or negative component will color information processing anytime the concept is involved. In other words, if we assume that the cat concept has a positive affective component in someone's mind, it is likely that perceiving a cat will trigger a positive affect. The evaluative priming task was designed to benefit from the cognitive architecture just described: It tries to capture the positive or negative component activated by a concept.

The evaluative priming task is a computerized task where participants must perform one thing: categorization. To capture the affective component, this task requires participants to categorize positive and negative words according to their affective valence (i.e., positive or negative; Fazio et al., 1986): these words are referred to as targets. Critically, targets will be preceded by words of different kinds, including the one for which researchers are interested in capturing the affective component. These preceding words are referred to as primes (see Figure 2.1). The researchers will also record the participants' response time, assuming that faster response times indicate that the primed concept eases the recognition of the affective dimension in the target and slower response times indicate that the primed concept interferes with the recognition of the affective dimension in the target. Usually, researchers assume that some of the concepts of which they are trying to capture the affective component are more likely to be either positive or affective. This assumption results in different types of trials: when a prime is followed by a target supposed to have the same affective valence, researchers talk about *compatible* trials, when it is followed by a target supposed to have the opposite affective valence, researchers talk about *incompatible* trials. Sometimes, researchers also include a type of prime which is not supposed to have an impact on the target word recognition (*control* trials, see Figure 2.1).

Figure 2.1*Time Course of the Evaluative Priming Task Trials*

Note. This figure is inspired from Fazio et al. (1986, Experiment 1). Participants must detect whether the target stimulus is positive or negative. Here, it is positive. The primes can either facilitate or interfere with the target's categorization.

According to Fazio et al. (1986), at the cognitive level, when the prime is processed, it is going to be activated in the cognitive system just like a turned-on light bulb. Because participants must categorize the target stimuli according to their affective component, the prime, coloring information processing, can either facilitate or interfere with participants' response (Fazio et al., 1986). Still assuming that cats possess a positive affective component,

we should expect that positive targets preceded by a cat will be categorized faster than negative ones. Fazio et al. (1986) provided a proof of concept of this protocol, showing the expected response latency for prime words participants reported as positive and negative. The following papers showed how useful the evaluative priming task could be.

It is important to remember that indirect evaluative measures were developed to deal with scenarios where participants would not want to disclose their answer on a direct evaluative measure. One sensitive area where people tend to keep their opinion for themselves is intergroup relationships: it is common for self-report answers collected in this context to be tainted with social desirability (Sigall & Page, 1971). People tend not to like being labeled as a racist. Following the seminal work on the evaluative priming task, in their paper, Fazio et al. (1995) computed an index indicating how much easier it was for participants to accurately identify a positive prime when it was preceded by the picture of a White American and a negative prime when it was preceded by the picture of an African American than the opposite. Interestingly, this index was more strongly correlated with a direct evaluative measure of racism (the Modern Racism Scale; McConahay, 1986) when participants had a low motivation to control their prejudice compared to when they were highly motivated. A possible interpretation of this result is that contrary to the Modern Racism Scale, participants do not (or do not feel the need to) exert a control on their performance on an evaluative priming task. Since then, this line of research has been prolific. The evaluative priming task produced results consistent with the capture of an affective component for novel stimuli (designed to be either positive or negative; Duckworth et al., 2002) and has even been used as a diagnostic tool to detect unrestrained eaters (Papies et al., 2009). Evidence for this task has been compiled into a meta-analysis (Herring et al., 2013) of 125 studies (with a total of 5,367 participants), showing an average effect size of $d_s = 0.37$,

95%CI[0.31; 0.43] indicating a small-to-medium (and reliable) difference between the different types of trial of the evaluative priming task.⁴

With the evaluative priming task, it seems that social psychologists developed a tool with the potential of limiting the problem related to the fact that people's answer to a direct evaluative measure could be influenced by motivation other than accuracy. Quickly, the evaluative priming task was followed by a second task—the IAT—which contributed in laying foundations of the implicit social cognition subfield (Greenwald & Banaji, 1995).

The Implicit Association Test

At the cognitive level, the theoretical ideas behind the IAT are different from the one found with the Evaluative Priming Task (Fazio et al., 1986). Rather than capturing the affective reaction resulting from the activation of a mental concept, the IAT captures the relative strength of the association between several mental concepts (Greenwald et al., 1998). To provide an intuitive example of what it means for concepts to be associated with each other, most people would agree with the idea that fishes and boats are somewhat more related to each other than fishes and books. In the theoretical framework in which the IAT operates, the question of the strength of the relationship between concepts is critical (Greenwald et al., 2002).

At the cognitive level, Greenwald et al. (2002) assume that concepts are stored in an associative network in which they are linked to one another. The consequence of an association between concepts is that the activation of one concept will result in the activation of the concepts it is associated with. Once a concept is activated, it is more likely that it gets used in information processing. Thinking about fishes should result in a higher likelihood to

⁴ In the present dissertation, we will adopt Lakens's (2013) recommendation and use the symbol d_s to report a between-subject Cohen's d and d_z to report a within-subject Cohen's d .

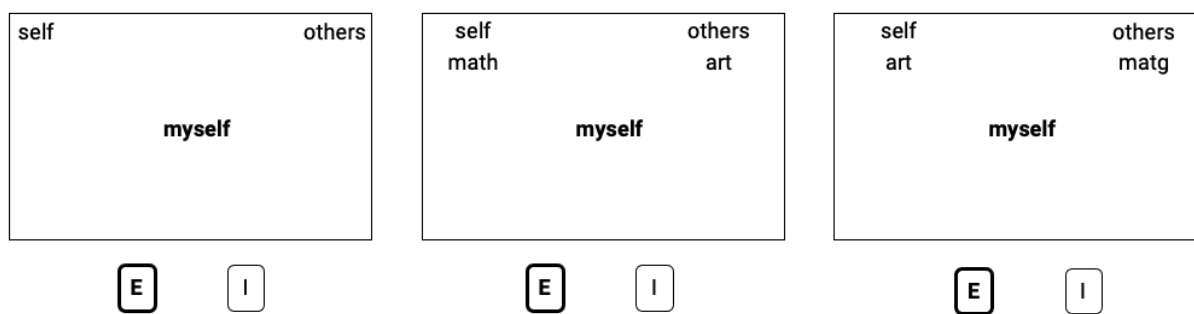
think about boats. In this associative network, some of the associations are particularly relevant for human behavior. The associations between the self and other concepts, for example, will drive people in certain directions (Greenwald et al., 2002). In this context, the IAT is a tool that can be used to know how two concepts differ regarding their associations to other concepts. For example, it can be used to know how much more someone associate the concept of self with math rather than art compared to the concept of others (Nosek et al., 2002).

To do so, just like the Evaluative Priming procedure, the IAT relies on response latencies in a computerized categorization task. Participants must categorize several target words according to the categories they belong to. Taking the example of an IAT aiming at assessing the relative strength of association between the self, the others, the math, and the art concepts, participants will have to categorize words like “myself”, “them”, “geometry”, or “poetry” according to the “self”, “others”, “math”, and “art” categories (Nosek et al., 2002). The special thing of the IAT which makes it possible for researchers to assess the relative strength of the relationship between concepts is the configuration in which one must perform the categorization task. In the IAT, participants must categorize the target words according to category labels which are displayed on the top corners of a computer screen (see Figure 2.2). If the target word belongs to a category displayed on the left corner, they must press a left response key (e.g., the letter “E” on a keyboard), if it belongs to a category on the right corner, they have to press a right response key (e.g., the letter “I” on a keyboard). Critically, the categories displayed in the upper corner will change during the IAT task. Usually, the first blocks involve categorizing targets along one type of categories (e.g., self vs. others or math vs. art). Critical blocks, however, involve participants categorizing targets along two types of categories at the same time (see Figure 2.2). Because an IAT is structured in

different blocks in which the configuration between top corner categories changes, researchers can assess whether it is easier to use the same response key to categorize self- and math-related words (along with others- and art-related words) compared to self- and art-related words (along with others- and pleasant-related words; see Figure 2.3).

Figure 2.2

Example of Configuration of the Categories in Different Blocks of an IAT



Note. Top part of the figure represents the onscreen information. Bottom part represents the keys that participant can press. Emphasized key is the correct one. Assuming that the stronger associations in memory are self-math and others-art, the block types from left to right are the practice block, the compatible block, and the incompatible block.

Figure 2.3

Block Structure for a Self-Others/Math-Art IAT

	Category label	Target word	Category label
Configuration 1	self		other
- Practice block	E	myself	I
	E	them	I
Configuration 2	math		art
- Practice block	E	geometry	I
	E	poetry	I
Configuration 3	self		other
- Practice block	math		art
- Test block	E	myself	I
	E	them	I
	E	geometry	I
	E	poetry	I
Configuration 4	art		math
- Practice block	E	geometry	I
	E	poetry	I
Configuration 5	self		other
- Practice block	art		math
- Test block	E	myself	I
	E	them	I
	E	geometry	I
	E	poetry	I

Note. This figure was adapted from Greenwald and Farnham (2000). Black keyboard keys represent correct response to the type of target in a specific configuration.

Based on these latencies during the task, researchers can compute indexes of the relative ease to perform the task in a configuration compared to another (Greenwald et al., 2003), the general idea remaining that a faster time to perform the categorization task indicates that the task is easier for the participants (see Richetin et al., 2015). In our example, a response time lower (i.e., participants faster) in the critical block where participants categorize the self and math with the same response key compared to the block where participants must use different response keys would be indicative of a stronger association between the self concept and math and others and art than the opposite (but see Blanton & Jaccard, 2006; Popa-Roch & Delmas, 2011; Rothermund & Wentura, 2004).

By focusing on the relative strength of associations of Black-White / pleasant-unpleasant or female-male / math-art, for example, the IAT has been used to investigate societal issues such as intergroup prejudice, or the gender gap in preference for math (Greenwald et al., 1998; Nosek et al., 2002). Indeed, under certain circumstances, the relative strength of association between concepts captured by the IAT is assumed to drive people's behavior (Greenwald & Banaji, 1995) and evidence that the IAT score can predict behavior is found in the literature (Greenwald et al., 2009; but see Oswald et al., 2013). A common idea present in the literature is that the IAT is better suited to predict spontaneous behaviors (e.g., Asendorpf et al., 2002; Dovidio et al., 2002). The idea that these indirect measures and spontaneity were somewhat associated shaped the research organized around the IAT, the field of implicit social cognition.

At the time these measures became popular, they offered a solution to some of the problems with self-report traditionally used in the attitude literature. By relying on the advances in cognitive psychology, researchers were able to build a new generation of tools to investigate human behavior. In this section, we introduced the measures before introducing the cognitive literature that inspired them. We will now turn to the literature on automaticity

and implicit memory, that is, the cognitive which inspired the evaluative priming task and the IAT. With regards to the question that we are addressing in the present dissertation, this cognitive literature plays a central role, because some predictions regarding the role of repeated experience can be traced to this literature.

The Theoretical Origins of Implicit Social Cognition

These two tasks (i.e., the evaluative priming task and the IAT; Fazio et al., 1986; Greenwald et al., 1998), inspired by cognitive psychology research, established the subfield of implicit social cognition (Gawronski et al., 2020). In a nutshell, implicit social cognition observed both methodological and theoretical developments, both converging toward a specific goal: accessing an otherwise inaccessible part of the human being. The theories elaborated in this context tried to account for the evaluative responses used by implicit social cognition researchers. In the implicit social cognition subfield, these evaluative responses are often categorized as either implicit or explicit, which we map onto the terms of direct and indirect evaluation to avoid the theoretical confusion that the other terms can generate (Corneille & Hütter, 2020).

Beside the evaluative priming task and the IAT (Fazio et al., 1986; Greenwald et al., 1998), a wide range of tools were developed in the implicit social cognition field and now falls under the umbrella of indirect evaluative responses. Nowadays, the variety of measures researchers can use is quite substantial: the Go/No-Go Association Task (Nosek & Banaji, 2001), the Shooter Bias Task (Correll et al., 2007), the Extrinsic Affective Simon Task (De Houwer, 2003), the Affective Misattribution Procedure (AMP; Payne et al., 2005), the Stereotype Misperception Task (SMT; Krieglmeier & Sherman, 2012), the Implicit Relational Assessment Procedure (Barnes-Holmes et al., 2010), or the Relation Responding

Task (De Houwer et al., 2015) or Approach-Avoidance Task (Chen & Bargh, 1999; Rougier et al., 2018; for a recent review on these measures, see Gawronski & Hahn, 2019).

As mentioned above, implicit social cognition is not only characterized by the wide range of paradigms it uses. Theories accounting for what these paradigms really capture are numerous (e.g., De Houwer, 2009; Gawronski & Bodenhausen, 2006; Rydell & McConnell, 2006). These theories usually involve theoretical concepts such as *implicitness*, *unconsciousness*, or *automaticity* which are typical of implicit social cognition. While a reader learning about the implicit social cognition literature could rapidly feel overwhelmed, it is possible to trace the origin of implicit social cognition to two theoretical perspectives which participated in the adoption of this vocabulary. In order to provide a clear overview of the landscape in which this manuscript takes place, we will discuss these theoretical roots which served as a foundation for implicit social cognition. Critically, these two roots are closely linked to the two tasks we just reviewed (i.e., the evaluative priming task and the IAT; Fazio et al., 1986; Greenwald et al., 1998).

The First Root of Implicit Social Cognition: Automaticity

In a recent review of the implicit social cognition subfield, Payne and Gawronski (2010) identified research related to selective attention and short-term memory as one of the foundations for implicit social cognition (e.g., Posner et al., 1976; Treisman, 1969). In short, this cognitive psychology literature focuses on the differences between automatic and controlled processes. Because the automaticity concept encompasses different notions (see Bargh, 1994; Moors & De Houwer, 2006), let us define what the automatic processes' characteristics are in this line of research. Research on selective attention and short-term memory was specifically dedicated to processes which need little attention, are unlimited in capacity, and are difficult to suppress. These processes differ from non-automatic ones which are defined as demanding in attention, limited in capacity, and voluntarily initiated (Payne &

Gawronski, 2010). Note that, contrary to later work on automaticity (e.g., Bargh, 1994), consciousness was not mentioned as a key distinctive feature between automatic and non-automatic processes. This line of research is the one that inspired Fazio et al.'s (1986) work.

The cognitive psychology literature interested in selective attention and short-term memory inspired Fazio et al. to build a task able to capture something appearing automatically as a result of a stimulus presentation (1986). According to Fazio et al., in terms of expression, “The key feature of such automatic activation [is its] inescapability” (1986, p. 229). The mere presentation of a stimulus can trigger an evaluative response which is captured by the task (through response time). This evaluative response, emerging automatically and captured by the evaluative priming task, was, for Fazio et al., the result of “a previously well-learned set of associations or responses” (Fazio et al., 1986, p. 229). To sum up, the first root of implicit social cognition assumed the existence of previously well-learned set of associations that researchers could capture through their uncontrollable automatic activation. With regard to this manuscript's question, that is, the role of repeated experience, it is important to note that we can find the notion of “well-learned set of associations” at the origin of the implicit social cognition subfield.

Now regarding the *implicitness* notion which is omnipresent in the implicit social cognition literature, *implicitness* sometimes characterizes the automaticity of the evaluative responses which is captured (Corneille & Hütter, 2020). This automaticity, captured by task such as the evaluative priming task (Fazio et al., 1986), was the first root of implicit social cognition. However, automaticity is not the only foundation of this subfield of social psychology; Greenwald et al.'s (1998) inspiration for the IAT came from a literature different than the one on automaticity.

The Second Root of Implicit Social Cognition: Awareness

The second line of research which inspired implicit social cognition is the one dedicated to implicit memory. It originates from cognitive psychology research and focuses on the impact of past experiences on present behavior, without a conscious recollection of the past experiences (e.g., Jacoby & Dallas, 1981; Warrington & Weiskrantz, 1968). In other words, this line of research investigated how something someone lived, but does not remember, influences how they behaves. An illustration of this idea is reported in *The Oxford Companion to the Mind* (Gregory & Zangwill, 1987, p. 149). The anecdote involved Claparède, a Swiss neurologist, and one of his patients, suffering from amnesia. Because of this condition, Claparède's patient could not remember her recent past; she was not even able to remember Claparède. Every day, Claparède was doomed to introduce himself to the patient, as if it was the first time they met. One day, Claparède hid a pin in his hand and shook the patient's hand, which resulted in a very unpleasant handshake for the patient. The next day, for no reason she was able to formulate, the patient refused to shake Claparède's hand. Far away from this entertaining example, research on implicit memory usually involves paradigm where a participant have to learn a set of stimuli, and then has to perform a test related to these stimuli, but does not require them to actually remember anything from the learning phase (Jacoby & Dallas, 1981).

This idea led Greenwald and Banaji to characterize implicit attitudes as “introspectively unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feeling, thought, or action toward social objects” (1995, p. 8). These attitudes would differ from explicit ones whose determinants can be identified and accessed. To capture these implicit attitudes, something which could not be done through introspection, Greenwald et al. (1998) developed a task where participants had to interact with a stimulus but did not have to recollect information regarding whether they like it: The

IAT. This second root clearly differs from Fazio et al.'s (1986) because its characteristics is the (un)awareness of the processes affecting a response.

Because of this definition, Greenwald and Banaji (1995) assumed two coexisting attitudes toward an object, the explicit and implicit ones. Advances in research with this conceptualization of implicit attitudes were mainly driven by methodological work with the IAT (Greenwald et al., 1998) and were notably fostered by research on its predictive validity. Among others, this literature shows how the IAT could uniquely predict certain type of behavior (for example, see Kurdi et al., 2018; Perugini et al., 2010), with for example, the IAT predicting spontaneous behavior and direct measures predicting controlled behavior (Asendorpf et al., 2002). These kinds of data were coherent with the idea of two coexisting attitudes in one's mind, one being implicit.

By reviewing the two procedures and the cognitive work that inspired them, we did not want the reader to focus on the differences between them, but the similarities. Importantly for this manuscript, the second root of implicit social cognition meets the first one in the importance it attaches to past experience. Indeed, its role is once again highlighted, as the initial research on IAT characterizes the construct captured by the IAT as the results of many traces of past experience, now introspectively inaccessible.

The Current Landscape of Implicit Social Cognition

As reviewed, the origin of implicit social cognition is diverse. Two theoretical roots can be identified as inspiring this field. One has to do with the uncontrability of evaluative responses and a second has to do their now unidentified origin. This multifaceted origin can be retrieved in the *implicit* term. Unfortunately, these are not the only concepts that the *implicit* term captures. In the current state of the affair, the *implicitness* of implicit social cognition can refer to, at least, five different properties of the measures, the cognitive

processes underlying them, or both (Corneille & Hütter, 2020). Despite the theoretical confusion generated by this term, during decades following the birth of implicit social cognition, researchers have used these tools to investigate a wide range of topics (e.g., political behaviors, substance use; Carraro et al., 2010; Payne et al., 2008). Discussions do exist regarding the features that the indirect evaluative responses actually possess (e.g., De Houwer, 2006a) and a certain consensus exist among researchers regarding some of the property of what is captured by the indirect evaluative responses. For example, nowadays, the idea that the indirect measures capture something outside of the sphere of awareness is now, at best, controversial (Corneille & Hütter, 2020; A. Hahn & Goedderz, 2020; but see, for example, Spatola & Wudarczyk, 2020). Still, some researchers consider that, at this time, theorization in implicit social cognition has still a lot of room for improvement (see Corneille & Hütter, 2020; Kurdi et al., 2018).

In this manuscript, we will limit our analysis to the question of the acquisition of indirect evaluative responses: How specific patterns of response on an IAT or other indirect evaluative response appear. Note that we will use the terms of evaluative learning and acquisition of evaluative responses interchangeably. We think that providing an overview of the context in which these measures were developed will help the readers understand how the assumptions regarding the role of repeated experience appeared. The idea that evaluative responses such as the Evaluative Priming Task or the IAT capture something which was well learned in the past will be critical regarding the present dissertation. While the contribution of repeated experience was highlighted in the original theoretical frameworks accompanying the Evaluative Priming Task and the IAT (Fazio, 1990; Greenwald & Banaji, 1995) these frameworks did not stay alone for long. The new millennium came alongside new theories accounting for how one would acquire an evaluative response and, as we will cover, these models also discuss the role of repeated experience. In the next section, we will review some

of the most influential models, namely, the System of Evaluation Model (SEM; McConnell & Rydell, 2014; Rydell & McConnell, 2006) the Associative-Propositional Model (APE; Gawronski & Bodenhausen, 2006, 2014a) and the single-process propositional model of evaluation (De Houwer, 2009, 2018).

Chapter 3

Theoretical Models of Evaluation

In the previous sections, we covered the context in which new kinds of evaluative measures such as the evaluative priming task or the IAT emerged. Even if we will not necessarily use these paradigms during the present dissertation (we will mainly focus on the IAT and on the Affect Misattribution Procedure; AMP; Payne et al., 2005), the evaluative priming task and the IAT had great theoretical importance in implicit social cognition. In the present dissertation, we are primarily interested in how people acquire specific behavioral responses on such measures. Earlier in the introduction, we discussed why we decided to adopt De Houwer et al.'s (2013) meta-theoretical framework. While this meta-theoretical model is relevant regarding the vocabulary we will use, it tells us little regarding the mental mechanisms which could account for indirect evaluative response acquisition. We will fill that gap by reviewing some of the modern models of evaluation acquisition. As it would substantially increase the length of an already long introduction section, we will not review in depth every model the implicit social cognition subfield has observed (see Payne & Gawronski, 2010), but stick to the ones which will be relevant regarding how people acquire indirect evaluative responses. Among others, the models we will not review include the MODE model (Fazio, 1990; Fazio & Towles-Schwen, 1999) or the Meta-Cognitive Model (MCM; Petty et al., 2007; Petty & Briñol, 2006). The reason why we decided to focus on the SEM (McConnell & Rydell, 2014; Rydell & McConnell, 2006), the APE model (Gawronski & Bodenhausen, 2006, 2014a), and the single-process propositional account (De Houwer, 2009, 2018) instead is because the cognitive mechanisms they discuss offer clear predictions regarding the contribution of actual experience in evaluation learning.

Before describing these three models, let us first describe a class of theoretical models that became extremely popular: dual models. The core assumption of dual models is that the humankind can process its environment in two different ways. One of the ways is fast, requires low effort, and is involved in intuition; The second is slow, effort consuming, and is involved in more deliberative thinking (Kahneman, 2011). This class of models has been extremely popular in the domain of evaluation processes (e.g., Smith & DeCoster, 2000; Strack & Deutsch, 2004). We will briefly discuss the two processes before addressing how they were integrated in the several theoretical model of evaluation. Note that, when reviewing these processes, we will make the distinction between operating principles and operating conditions of a mechanism (e.g., J. W. Sherman et al., 2014). Operating principles refers to *how* processes work. Describing that there is a mechanism strengthening the association between two concepts in an associative system is an operating principle. Operating conditions refers to *when* a process work. Describing that there is a process which operates in an efficient manner is an operating condition.

The Two Types of Processes in Dual Models

Associative Processes

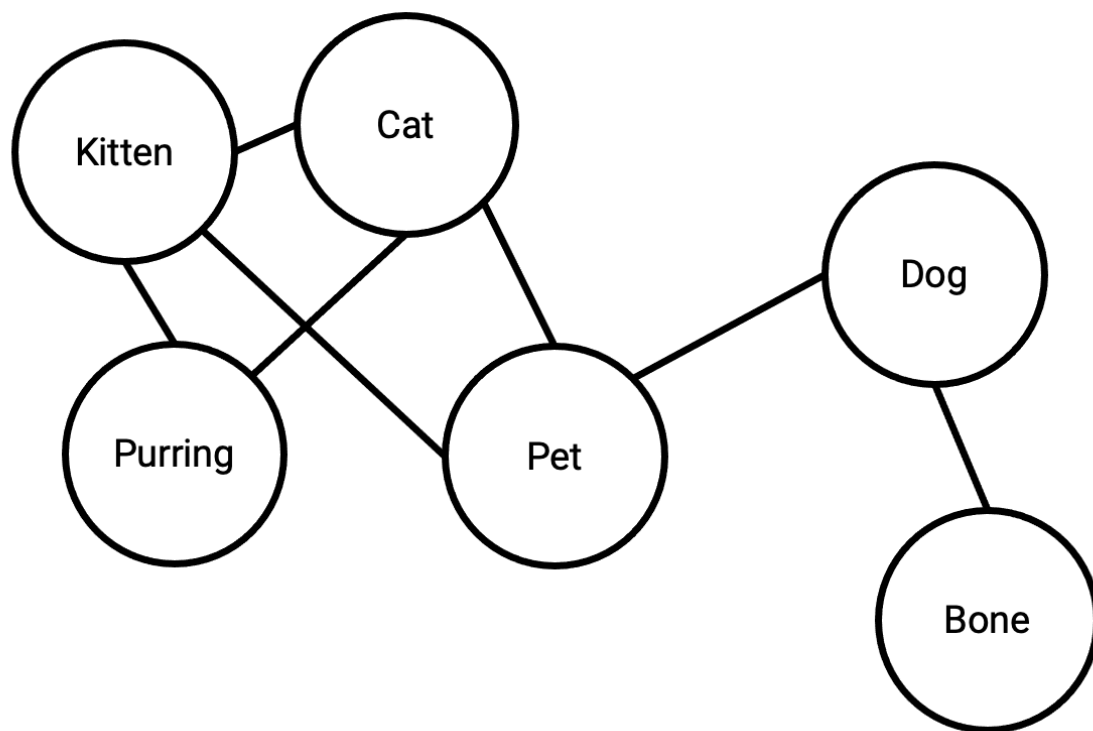
Associative processes are one of the two types of processes usually formalized in dual models of evaluation. They differ from the other processes—the propositional ones—in the way they are assumed to operate: They are the fast low-effort mechanisms involved in intuitive thinking. Their use is often seen as operating automatically, that is to say in an unconscious, efficient, unintentional, and uncontrollable fashion (for a discussion on automaticity, see Bargh, 1994; Moors & De Houwer, 2006). A large part of the literature dedicated to evaluative responses has used these processes to explain phenomena linked to indirect evaluation (acquisition, expression; Hughes et al., 2011). In this section, we will

focus on how associative processes are assumed to operate. In other words, we will focus on the operating principles of associative processes.

In terms of cognitive representations, these processes often assume that people store information in an associative network. In this network, one can find nodes, each node representing a mental concept, and nodes can be associated with each other (see Figure 3.1). The structure of this associative network will depend on the associations between the mental concepts and will determine the indirect evaluation of stimuli. The principle used to create associations within the network is simple and can be summarized in one sentence: “What fire together, wire together” (Hebb, 1949). Associations between concepts will appear when they co-occur on the basis of a spatial and temporal contiguity (Gawronski & Bodenhausen, 2006). That is, if I am exposed to purring any time I see a kitten, an association will appear between the concepts of “purring” and “kitten”. The learning of new association is often considered slow and gradually happening over time (Smith & DeCoster, 2000).

Figure 3.1

Schematic Illustration of How Information Is Stored in an Associative Network



Note. Each node represents a concept. Associative processes assume that links between nodes usually emerged as a result of the perception of co-occurrence between concepts.

For some evaluations theorists (Gawronski & Bodenhausen, 2006, 2014a; McConnell & Rydell, 2014; Strack & Deutsch, 2004), this way to store and organize information will affect indirect evaluations. In an associative network, the activation of a concept will depend on the degree of similarity between the said concept and the subjective experience of the environment. Put differently, if I have a cat in front of me, it is likely that the concept *cat* will be activated in my associative network. Critically, once the cat concept is activated, so will be the concepts associated with it—a phenomenon called activation spreading. So far, we talked about associations between concepts, without going in depth with the meaning of these associations. Associative processes posit that these associations do not carry any property

regarding the nature of the relation between the concepts. An association is simply an excitation link between a mental concept and another, which will make it likely that a concept will be activated once the other is (Shanks, 2007).

Dual models of evaluation usually describe activation spreading as what will cause an indirect evaluation. Let's take again the cat concept. When perceiving a cat, it is likely that the concepts of purring and pet will be activated because of activation spreading (see Figure 3.1; p. 48). Even if, initially, someone has a neutral reaction toward a cat, the activation spreading phenomenon could give rise to an overall positive reaction if we assume that the purring and pet concepts carry a positive valence. Evaluative procedures such as the evaluative priming task are designed to capture this affect (Fazio et al., 1986).

An important feature of these associative networks is that the whole network is not activated at once; rather, concepts are reactivated in a pattern completion fashion (Smith & DeCoster, 2000). This operating principle explains how the same stimulus can sometimes trigger different evaluation according to the context in which it is presented (e.g., Wittenbrink et al., 2001).

Critically, at no time one needs to endorse an association for the association to influence an indirect evaluation (Gawronski & Bodenhausen, 2006). Social psychologists have tackled issues with broader societal implication than the impact of perceiving a cat. If, any time I am exposed to a Black American in the media, violence is involved, an association between Blacks and violence will emerge. I do not need to agree with the sentence "Black people are violent" for an indirect evaluation congruent with this proposition to emerge (Devine, 1989). For this reason, it has sometimes been argued that indirect evaluations capture cultural knowledge instead of something personal (Arkes & Tetlock, 2004).

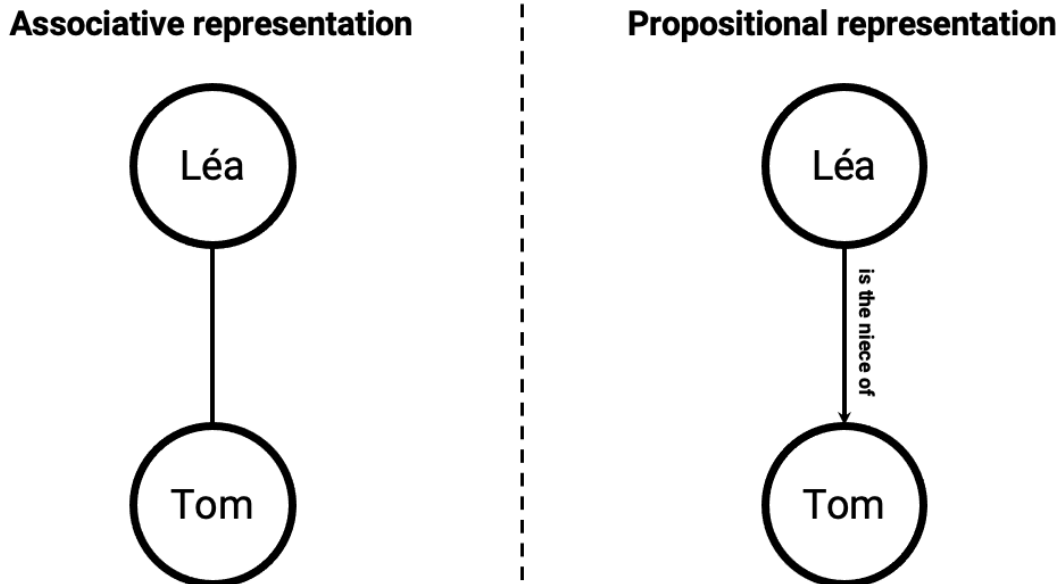
The operating principles behind associative networks have been used to explain when an indirect evaluation will emerge, but also to explain discrepancy between direct and indirect evaluations. Indeed, dual models of evaluation usually suggest that direct evaluations are governed by different processes than indirect evaluations. In the next section, we will review the processes mainly responsible for direct evaluations: propositional processes.

Propositional Processes

While indirect evaluations are theoretically mainly driven by associative processes, dual models of evaluation suggest that direct evaluations are mainly driven by propositional processes. In comparison to the first type, the second type of processes is often seen as slower, requiring effort, and involved in reasoning (Kahneman, 2011). In this section, we will discuss the operating principles behind propositional processes of evaluation.

In terms of cognitive representation, a critical point on which propositional processes differ from associative ones is the type of relation they operationalize between concepts (Strack & Deutsch, 2004). While the associative network allows meaningless association between concepts, propositional processes operate on the basis of qualified association between concepts (i.e., proposition). Assuming a relationship between two people, propositional processes can describe a level of specificity such as “Léa is Tom’s niece” (see Figure 3.2). Associative processes would have simply translated this relationship by a mere excitatory link between Léa and Tom.

A critical implication of this structure between concepts is that a truth value can be assessed for a proposition. Indeed, if we consider the two following propositions “Léa is Tom’s niece” and “Léa is Tom’s aunt”, one can infer that one of the two propositions is false (assuming a traditional genealogy); Associative processes would have translated these two associations into the same excitatory link between Tom and Léa.

Figure 3.2*Representations of Associative and Propositional Relationships*

Note. In the left panel (associative representation), the association represents a mere excitation link. In the right panel (propositional representation), the associations represent symbolic relationship between mental concepts.

Association in associative processes are assumed to emerge following the perception of co-occurrences. In propositional processes, Strack and Deutsch (2004) suggest that propositions can either come “from the outside world” or “from the inside”. First, because human beings communicate and share information with each other, propositions can be learned socially. They can also emerge from one’s own reasoning. Because propositions have truth values, logic rules apply. If I know that Paul is Léa’s father and if I know that he also is Tom’s brother, I can infer that Léa is Tom’s niece and Tom is Léa’s uncle. Some authors refer to these propositions emerging from other propositions as inferences (Van Dessel et al.,

2018a). Note that a proposition can also emerge from an activation spreading in an associative network. When an affective reaction emerges from the perception of a cat (because of the “purring” concepts it is linked to, propositions such as “I like cats” can emerge; but see McConnell & Rydell, 2014).

Regarding the storage of propositions, consensus appears to be less important than it is for associative processes. Some models suggest a dedicated system to deal with propositions (e.g., Rydell & McConnell, 2006), while others argue that both associative and propositional system can occur in the same information storage (see Gawronski & Bodenhausen, 2018).⁵ It has to be noted, however, that the question of how propositions are stored does not play an important role in the mechanisms involved in direct evaluation (unlike associative processes).

While associative processes describe spreading of activation as one of the mechanisms mediating evaluation, propositional processes describe another mechanism called proposition validation. When one replies to a self-reported question, they will either have to produce an answer or select one among a set of provided ones. To do so, dual models usually assume the individual will consider a subset of relevant propositions. This set of

⁵ Note that this theoretical point is a subject of debate for researchers. Gawronski and Bodenhausen (2018) argue that associative networks can theoretically produce propositions (but see Hummel, 2010). This involves a multi-layer connectionist model where a first layer would connect concept with each other and second one would capture the relationships between concepts. However, such structure could be problematic regarding the mere definition of associative processes. Indeed, if relational information can be represented in an associative network De Houwer (2018) argues that it falls under the propositional processes umbrella.

relevant propositions is what a respondent will use to produce a direct evaluation. In some circumstances, the set of considered propositions can contradict itself. For example, when an individual is asked whether they feel self-confident, they could consider the following set of propositions: “I don’t feel self-confident”, “Self-confidence is something valued”, “I want to be well perceived”. Dual models usually assume that these situations will lead to an aversive dissonance (see Festinger, 1968) and individuals will usually disregard one of the propositions or engage into elaboration in order to find a new proposition solving the conflict to avoid the dissonance. In our example, producing the answer “I feel self-confident” solves the conflict.

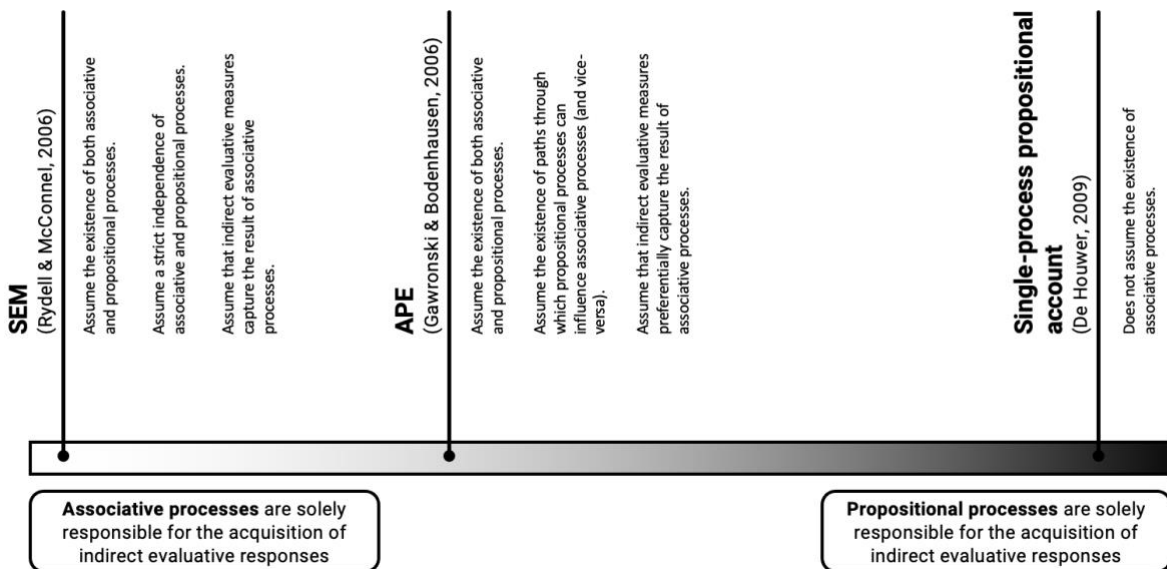
In most models of evaluation (e.g., Gawronski & Bodenhausen, 2006; Rydell & McConnell, 2006; Strack & Deutsch, 2004), the propositional processes are dedicated to the explanation of direct evaluative responses and only the explanation of direct evaluative responses. However, models such as De Houwer’s single-process propositional model (2018) also uses propositional processes to explain indirect evaluative responses. In such a model, the influence of a proposition during the production of an indirect evaluative response occurs on the basis of automatic similarity-based retrieval processes (see De Houwer, 2018). We will not use the present section to review such a mechanism but will describe it in the section dedicated to single-process propositional models.

Theoretical Models of Evaluation

Dual process models of evaluation use these two mechanisms—associative and propositional—to account for people’s evaluative responses. In this literature, the associative processes are mainly used to account for indirect evaluative measures (Hughes et al., 2011). Interestingly, these associative processes map relatively well on what Fazio et al. (1986) and Greenwald and Banaji (1995) described in their early work on a new generation of evaluative

measures. Remember, Fazio et al. described what their measure captured as the result of “previously well-learned set of associations or responses” (1986, p. 229) and Greenwald and Banaji assumed that the IAT was supposed to capture were the result of “introspectively unidentified (or inaccurately identified) traces of past experience” (1995, p. 8). The associative processes that we described in the previous section are not different from these original statements regarding the importance of repeated experience in the acquisition of indirect evaluative response.

Still, several models exist. These models differ regarding their operating principles—that is, how the different processes are going to work—but also regarding their operating conditions—that is, when the processes are going to work (see Gawronski & Bodenhausen, 2009; J. W. Sherman et al., 2014). Even if we described the two associative and propositional processes as independent from each other, theoretical models of evaluations sometimes assume complex interaction between associative and propositional processes which result in dual models quite different from each other (e.g., Gawronski & Bodenhausen, 2006, 2014a; McConnell & Rydell, 2014; Rydell & McConnell, 2006). In the next section, we will focus on two of these dual models, as well as another type of model: single process propositional model. Critically for the associative and propositional processes that we just introduced, the models on which we decided to focus can be placed along a continuum indicating how much indirect evaluation acquisition is specific to associative processes (see Figure 3.3). As we will see later, this property will be relevant regarding the predictions we can make of the role of repeated experience in indirect evaluation acquisition. For now, we will review the SEM (McConnell & Rydell, 2014; Rydell & McConnell, 2006), the APE model (Gawronski & Bodenhausen, 2006, 2014a), and the single-process propositional model (De Houwer, 2009).

Figure 3.3*Importance of Associative Processes in Indirect Evaluation*

Note. The different theoretical model reported here assumed different level of specificity of indirect evaluation to associative processes.

APE = Associative Propositional Evaluation, SEM = Systems of Evaluation Model.

The SEM

As mentioned earlier, the SEM (sometimes referred to as the systems of reasoning account) belongs to the general category of theoretical models that explains the human psychology with two distinct mechanisms (McConnell & Rydell, 2014; Rydell & McConnell, 2006). In terms of operating principles, the SEM suggests that two independent systems operate and are responsible for evaluative responses: The first one is the associative system and accounts for indirect evaluative responses, the second one is the rule-based system (hereafter referred to as the propositional system for the purpose of consistency between models) and accounts for direct evaluative responses.

As stated when we provided an overview of the theoretical models on which we will focus, the SEM can be placed along a continuum describing the extent to which indirect evaluations are specific of the associative system. On this continuum, the SEM is located at one of the extrema. The SEM assumes that the associative and propositional processes occur in two independent systems. One of the two systems works according to the associative processes described above, the second one according to propositional processes. In their seminal paper, Rydell and McConnell (2006) described the associative system as the slow-learning system which learn through repeated paired associations based on similarity and contiguity over time, and the propositional system as the fast-learning system which relies on higher order cognitive level of cognitive processing to handle symbolic or verbal representations. The associative system is responsible for the automatic responses toward stimuli while the propositional system is responsible for deliberative behavior. Such description maps onto a sole impact of associative processes on indirect evaluation and a sole impact of propositional processes on direct evaluation. To understand why Rydell and McConnell framed the SEM like this, we will focus on the prototypical experiment backing up the model.

The prototypical data supporting the two independent associative and propositional systems come from an experiment in which participants learned direct and indirect evaluative responses of opposite valence at the same time (Rydell et al., 2006). In this experiment, participants were exposed to information about a character named Bob and they could infer from this information that Bob was a bad (good) person.⁶ Simultaneously, participants were

⁶ Participants were not directly exposed to negative information about Bob. Rather, they were first exposed to a behavior and had to guess whether the behavior was characteristic of Bob.

subliminally repeatedly primed with positive (negative) words. The results were in line with what would have predicted the SEM—assuming that participants could form associations, but not propositions, based on the subliminal primes. Rydell et al. (2006) observed that participant's direct and indirect evaluations were consistent with, respectively, inferred and subliminal information. When participants read negative information, their evaluation of Bob was negative on several direct evaluative response measures (e.g., scale evaluating how much Bob was likable, feeling thermometer). When participants were subliminally primed with positive words, they were faster for the blocks where Bob and positive attributes were associated with the same response key than when they were not during an IAT, hence exhibiting a positive indirect evaluative response. The SEM accounts for the results with the two associative and propositional processes operating simultaneously. Each of the processes learned an evaluative response in a dedicated system. Here, the associative processes would have integrated the association between Bob and the positive prime in an automatic fashion—the prime being subliminally presented, and the propositional processes would have produced a positive direct evaluation using the positive information from the learning phase. This experiment can be considered as a clear illustration of the separation between the two associative and propositional processes. Note, however, that while the “Bob's paradigm” has been used several times to support the SEM (e.g., McConnell et al., 2008) a preregistered replication of Rydell et al.'s (2006) study recently failed at replicating the results of this experiment (Heycke et al., 2018).

After their response, participants received feedback indicating whether the behavior was indeed characteristic of Bob. Thus, participants had to elaborate in order to conclude that Bob is a bad person.

In the end, because of this strict distinction between the associative and propositional systems, few methods exist for one to acquire an indirect evaluative response. Because the SEM suggests that the associative processes occurring in the associative system are solely responsible for indirect evaluation, actually perceiving co-occurrence repeatedly appears to be the only way a person could develop a specific response on an IAT or an Evaluative Priming Task. The role of actual experience appears to be critical here. Compared to the other models we will focus on, because of this strict dissociation between the associative system and the propositional system, there is no room for propositional processes to influence an indirect evaluation⁷. This strict distinction led some authors to consider the SEM “the most radically dual of the dual-learning models” (Corneille et al., 2019, p. 178). Indeed, among dual models of evaluation, some allow for more permeability between the associative and propositional learning mechanisms.

The APE Model

In terms of strict dissociation between the propositional and associative processes, the APE model is a less radical cousin of the SEM (Corneille & Stahl, 2019; Gawronski & Bodenhausen, 2006, 2014a). This model is one of the most influential models of evaluation (with more than 2300 citations for the 2006 seminal paper according to a Google Scholar query) and it also uses the two processes described above to account for direct and indirect evaluations. However, unlike the SEM, the APE model assumes complex interaction between

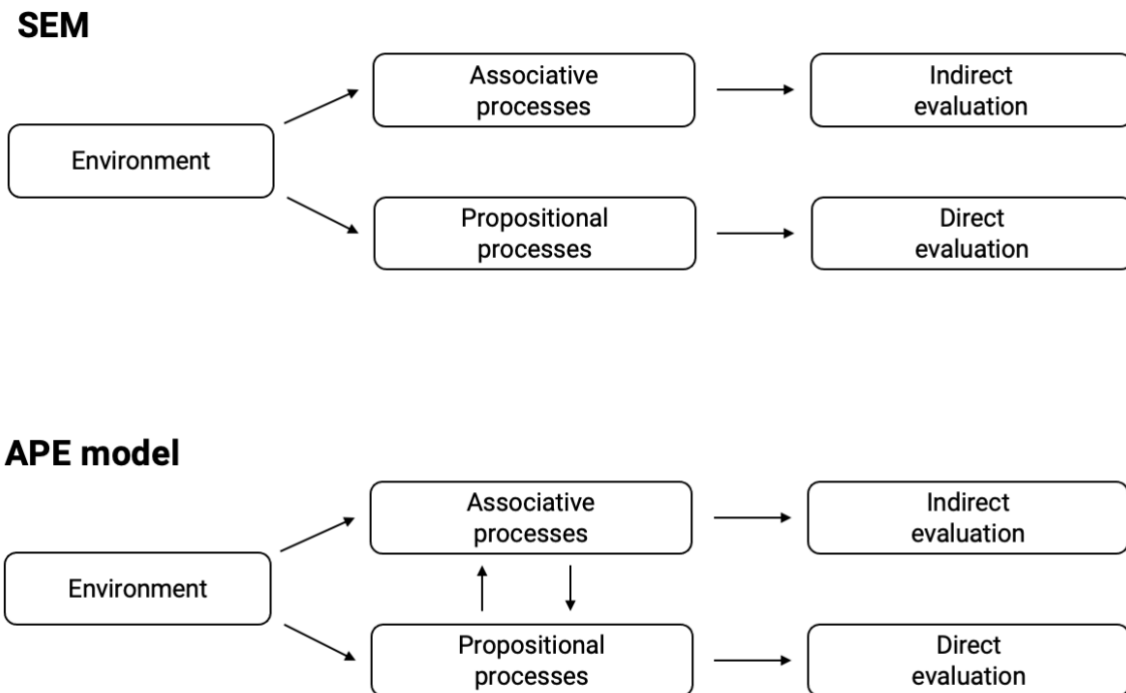
⁷ There are some rare instances in which the SEM acknowledges associative processes influencing direct evaluations and the propositional processes influencing the indirect evaluations. This happens mostly when one does not possess any “information” in the system that the evaluative response measure preferentially captures (e.g., no associations with a stimulus presented in an evaluative priming task).

the two processes rather than a strict dissociation. According to its authors, the APE model's main contribution to the literature lies in this complex interaction, as it allows the APE model to account for a wider range of data in the literature (Gawronski & Bodenhausen, 2006, 2014a). To review this model, we will first discuss the similarities with the SEM, and then, cover the divergences.

Like the SEM, the APE model belongs to the category of dual process models of evaluation. There is no difference between the APE model and the SEM in that the associative processes describe how associations between mental concepts generate indirect evaluation and propositional processes describe how an individual generates an evaluation on the basis of a set of consistent propositions. However, unlike the SEM, the APE model description of associative and propositional processes does not imply two distinct memory systems for the two processes to operate (for a recent discussion on this topic; see Gawronski & Bodenhausen, 2018). A critical feature of the APE model is the existence of indirect routes through which one can acquire evaluative responses (see Figure 3.4). Indeed, in the APE model people can acquire a direct evaluation because of the associative processes and an indirect evaluation because of propositional processes.

Figure 3.4

Difference of Operating Principles between the SEM and the APE Model



Note. Whereas the SEM formalizes a strict dissociation between associative and propositional processes, the APE model does not. APE = Associative-Propositional Evaluation, SEM = Systems of Evaluation Model

In terms of operating principles, the APE model suggests the existence of two phenomena allowing for the indirect routes of evaluation acquisition. The first one is the ability for propositions to emerge because of activation in associative network. Indeed, if we go back to our example where cats are associated with positive concepts in the associative network, the mere perception of a cat and the phenomenon of activation spreading could produce a proposition such as “I like cats” in the mind of the perceiver. The second phenomenon accounting for indirect routes is the ability of reasoning to form an associative link between concepts. Repeatedly thinking about two concepts could result in an associative

link between the concepts in a way similar to the encoding of repeated co-occurrences.

Because of these indirect routes, the APE model is different from the SEM in the number of evaluation acquisition for which it can account. We will review in more detail the kinds of data these indirect routes allow the APE model to account for in the next chapter.

To understand the implications of these differences between the APE model and the SEM, we can focus on indirect evaluation acquisition through mere instructions. Some experimental paradigms used in implicit social cognition involve participants merely receiving instructions related to, for example, a repeated pairing between two stimuli (see De Houwer & Hughes, 2016). Readers familiar with the literature might recognize here the kind of data showing the acquisition of an evaluative response toward a stimulus as a result of the perception of repeated pairings of the said stimulus with a stimulus of positive or negative valence (De Houwer, 2007). In this kind of paradigm, participants do not actually experience a repeated exposition to the stimuli. If they did, every dual model would assume that an indirect evaluative response would emerge because of associative processes, but in the absence of the repeated experience, the predictions of the APE model and the SEM differ. The APE model assumes that an indirect evaluative response can emerge as long as propositional processes successfully built a link between concepts in the associative networks; the SEM assumes an indirect evaluative response cannot emerge without repeated experience. Experimental evidence shows that merely being exposed to the instructions—without actually experiencing any pairing—results in an indirect evaluation acquisition (Gast & De Houwer, 2012). These data are hardly explained by the SEM, as it postulates a strict distinction between propositional associative processes and propositional ones. In contrast, the existence of an indirect evaluation acquisition in the APE model fits with such experimental evidence.

Regarding the operating conditions of the associative and propositional processes, the APE model explicitly rejects the hypothesis that associative processes are entirely automatic processes while propositional processes are entirely controlled. In the words of Gawronski and Bodenhausen: “According to the APE model, there is no one-to-one mapping between operating principles and operating conditions, such that associative processes generally operate automatically, whereas propositional processes generally operate in a controlled fashion. Instead, both associative and propositional processes have automatic and controlled aspects.” (Gawronski & Bodenhausen, 2014b, p. 458). Nevertheless, in their review dedicated to associative learning, Corneille and Stahl (2019) report a high degree of overlap between automatic processes and the associative ones and controlled processes and the propositional ones.⁸

Finally, a critical question for the present dissertation is the role of repeated experience in indirect evaluative learning. The existence of indirect routes allows the APE model to account for empirical data where mere instructions (rather than repeated experience) results in an indirect evaluative response acquisition (e.g., Gast & De Houwer, 2012; Van Dessel et al., 2015). However, as reported in a Gawronski and Bodenhausen’s (2018) paper, evidence still suggests something special about actual experience when it comes to evaluation acquisition, as illustrated by instances in which repeated experience changes an evaluative response while mere instructions do not (Hu et al., 2017; see also Van Dessel et al., 2020).

⁸ The definition of the operating conditions of the associative and propositional processes in the APE model goes beyond the scope of the present manuscript. Readers interested in this topic will find in-depth discussions in Corneille and Stahl (2019) and Gawronski and Bodenhausen’s (2018) papers.

As illustrated in this section dedicated to the APE model, dual models of evaluation do not necessarily entail a strict dissociation between associative and propositional processes. The APE model, however, is not the only model describing how propositional processes can impact the acquisition of indirect evaluative response. Single-process propositional models, as the name indicates, describe a category of models that rely solely on propositional processes to explain the acquisition of evaluative responses, be they direct or indirect. In doing so, they reject the existence of associative processes. We will review these models in the next section.

Single Process Propositional Models

The models that we reviewed so far belong to the dual process category. This category has been very popular in implicit social cognition and, in fact, most theories use associative processes to account for indirect evaluation (Hughes et al., 2011). Still, alternatives exist and some authors suggest that a single process can account for both direct and indirect evaluation acquisition (e.g., De Houwer, 2009, 2018; Mitchell et al., 2009; Van Dessel et al., 2018a). In this section, we will review a single process propositional model of evaluation acquisition, focusing on De Houwer's (2018) formalization. This single process propositional model is distinguishable from dual models in several domains that we will discuss. The main difference can be summarized by the rejection of the hypothesis of the evaluative learning through associative processes. Because associative processes often encompass operating principles and conditions, single-process propositional models stand out from dual models in several ways.

In terms of operating principles, the single-process propositional models adopt proposition as the basis of information organization. Propositions, as discussed above, are different from mere associations as they contain information about the nature of the relation

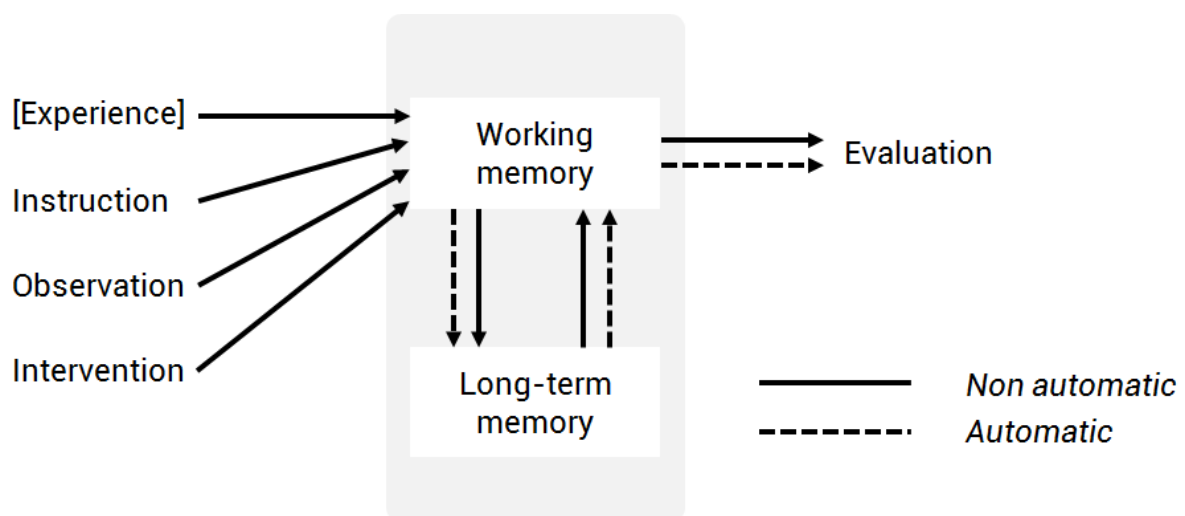
between the two mental concepts they pair. In other words, while associative processes describe that a repeated exposure to sentences like “Léa is Tom’s niece” or “Léa is Tom’s aunt” result in the emergence of the same mere association—that is, an excitatory link—between the Léa and Tom concepts, propositional processes would not (see Figure 3.2, p. 51). Along with a link between concepts, propositions come with the qualification for the relation and, therefore, with a truth value (i.e., a proposition can theoretically be considered true or false). In terms of evaluative learning, this operating principle results in indirect evaluation learning being moderated by relational information between concepts (a prediction associative processes do not offer without ad hoc explanations; see Gawronski & Bodenhausen, 2018). However, single-process propositional models also differ in the operating condition it assumes when it comes to evaluation acquisition.

Contrary to dual processes of evaluation, single-process propositional models often reject the assumption that learning can occur according under the operating conditions of automaticity. They usually assume that the formation of propositions (ultimately influencing evaluative responses) requires awareness, time, effort, and a goal to learn (De Houwer, 2018). As influential as the idea of an automatic learning of indirect evaluative response can be in the literature (e.g., Rydell et al., 2006), a recent implicit social cognition literature review suggests that the learning of indirect evaluative response is not as automatic as the literature often claims (Corneille & Stahl, 2019). Dual models assume that every single co-occurrence in the environment is registered in the associative system—which often results in the formation of evaluative responses. Single-process propositional models oppose to this assumption the argument that if every co-occurrence had to be registered, organisms would probably behave in erratic ways. The registered co-occurrences must be selected based on some criteria. De Houwer’s (2018) assumes that the co-occurrences an individual will register are the one that will help them optimize their behavior. The selection process, such as

other processes like problem solving, is assumed to be non-automatic (see also Mitchell et al., 2009). Note, however, that while single-process propositional models usually assume evaluative learning to be non-automatic, it does not assume that evaluative expression also is (see Figure 3.5).

Figure 3.5

Source of Evaluative Learning According to the Single-Process Propositional Model



Note. Adapted from De Houwer (2018). Bracketed terms have been edited from the original figure to better fit with this dissertation.

The single-process propositional model described in De Houwer (2018) suggests that propositions are stored in an episodic memory system. Such models usually store an event as an episodic trace (e.g., Hintzman, 1984; Versace et al., 2014) and episodic traces are usually described with the ability to be automatically reactivated on the basis of the similarity between the environment and the content of the episodic trace. Consequently, propositions can be reactivated as the result of a stimulus perception. Just like with associative processes, a reactivated trace is likely to influence behavior. Note that the reactivation of episodic traces

is very similar to the activation spreading in associative processes and, in fact, the episodic memory system encoding propositions could be formalized in an associative network (De Houwer, 2018). The difference between the two mechanisms in single-process propositional models and dual models lies in the content of the reactivated association/traces and their origin. Regarding their content, as discussed above, dual models assume a mere association between concepts. On the contrary, single-process propositional models suggest that links between concepts contain the nature of the relationship (see Figure 3.2, p. 51). Second, whereas dual models assume that the formation of association is mostly automatic, single-process propositional models reject this assumption (De Houwer, 2018).

Does it mean that single process propositional models assume no impact of the repetition of an experience on (indirect) evaluation acquisition? The answer is no. In a single process propositional model explaining how individuals acquire evaluative response through approach and avoidance behaviors, Van Dessel et al. (2018a) argue that the repetition of an experience can influence acquisition of an evaluative response. Remember, in single-process propositional models, encoding a proposition is a non-automatic process (which should be understood as awareness-requiring, time-consuming, effortful, and goal-dependent processes; De Houwer, 2018). Any factor likely to influence the formation of a proposition is likely to influence evaluation acquisition. It does not mean that the repetition of an experience will always moderate the acquisition of an evaluative response, but in certain situations, it can help. A real-life example which should help the reader understand the nuance is a Christopher Nolan and Michael Bay. It is likely that several watching of a Christopher Nolan movie will help the viewer understand it; On the other hand, a single watch of a Michael Bay movie is reasonably enough to understand the whole complexity of the plot, no matter how good it is. In some instances, actual experience should help the acquisition of an evaluative response.

In the end, single-process propositional models are extremely different from radical dual models such as Rydell and McConnel's SEM, as they reject the existence of associative processes (De Houwer, 2009, 2018; Van Dessel et al., 2018a). The main difference between the single process propositional models and the dual ones remain the rejection of automatic learning of evaluative response from the perspective of single-process propositional models. Finally, regarding the role of repeated experience, critical in the present manuscript, as reviewed, single process propositional models do suggest that actual experience could impact evaluation acquisition in a non-systematic manner.

Theoretical Models of Evaluation and the Role of Actual Experience

In the present dissertation, we are mainly interested in indirect evaluative responses. It is common for these measures to be merely referred to as implicit measures. Without additional explanations, this can be confusing (Corneille & Hütter, 2020). We hope that this short review of the models of implicit social cognition will help the readers navigate this literature. Here, what is important to know is that, historically, associative processes have been preferred to explain indirect-evaluative-response-related phenomena (Hughes et al., 2011). The relative importance of associative processes in indirect evaluation acquisition now varies across theories—up to a point where it is not involved at all, such as in single process propositional models (De Houwer, 2009, 2018; Van Dessel et al., 2018a). Reviewing in more details every model of indirect evaluation, their operating principles, and operating condition goes beyond the scope of the present manuscript. Here, we are mainly interested in one factor and its role assumed in evaluative response acquisition: repeated experience.

An important issue that we want to raise here is that, in the present dissertation, we will not be interested in comparing the different models of evaluation. Rather, we will use the predictions of the models to investigate the question of repeated experience. A consequence

of this approach is that our experiments will not be designed to maximize the differences between the prediction of the several models. Rather, we will explore different situations in which some of the models predict a contribution of repeated experience in evaluative learning. We want to stress that, in some of our experiments, some theoretical models (but not all of them) will stay agnostic regarding a contribution of repeated experience. We will make the effort to make it explicit how the results of our experiments fit with the existing literature. The goal of our short review of some of the models of evaluation was to better understand how the different models integrated some of the assumptions of the early days of implicit social cognition regarding the role of repeated experience.

Historically, associative processes are tied to the Hebb's rule of "what fires together, wires together" (Hebb, 1949) and, in this principle, lies the idea that repeated experience is critical when it comes to indirect evaluation acquisition. Every model we review comes with specific prediction related to repeated experience. It will not be surprising for the reader familiar with the implicit social cognition literature, as some of the paradigms used to investigate indirect evaluative response acquisition adopt some sort of repetition in their design. The use of these paradigms, as well as recent development in experimental procedures, will be the main focus of our next chapter.

So far, we barely mentioned the actual paradigms related to indirect evaluation acquisition. In the next section, we will review some of these paradigms. By doing so, we will see that the assumption that repeated experience is necessary for indirect evaluative learning has been influential in the implicit social cognition literature. Even if it has not always been explicitly formalized as is, a lot of paradigms used in implicit social cognition research rely on participants repeatedly experiencing something. We will first review some of the experience-based paradigms, including the approach-avoidance training paradigm, central in the present manuscript, and then, we will present instructions-based counterparts of these

repetition-based paradigm. In doing so, we will start discussing the core question of this manuscript: What is actually the added value of actual experience in indirect evaluative learning?

Chapter 4

Observing Evaluative Learning

In the present dissertation, so far, we focused on how some social psychologists investigated consistency in human behavior. We saw that they developed a special kind of evaluative measures which had the potential to overcome some of the limitations of self-reported measures. Along with these indirect evaluative measures, social psychologists developed theoretical models accounting for acquisition of these evaluative responses. However, so far, we barely mentioned how researchers were actually investigating evaluation acquisition. In the present chapter, we will review some of the paradigms used to study indirect evaluation acquisition and introduce the main paradigm that we will use in the present dissertation. Relevant to the question of the contribution of repeated experience to indirect evaluation acquisition, we will make the distinction between two types of paradigms: the paradigms based on the repeated-experience principle and the ones based on instructions. We will start by reviewing repeated-experience-based paradigms, before discussing their instructions-based counterparts.

Paradigms Based on Repeated Experience

In the current section, we will not conduct an exhaustive review of the paradigms used to make people acquire new indirect evaluative responses, as this would not align with the current scope of the manuscript. The paradigms we decided to review were chosen because of their relevance regarding the question that we are addressing in this manuscript and because of their importance in implicit social cognition research. We will review the Evaluative Conditioning paradigm and approach-avoidance training, a paradigm which will be the main focus of the present dissertation. Readers interested in a more complete view of

the paradigms aiming at changing indirect evaluation will find resources in Forscher, Lai, et al. (2019) or Lai et al. (2014). These papers describe paradigms fitting in implicit social cognition research, but that we decided not to review (e.g., being exposed to counterstereotypical exemplars; Dasgupta & Greenwald, 2001), as well as paradigms inspired by other research tradition (e.g., self-affirmation or implementation intention; Rudman et al., 2007; Stewart & Payne, 2008).

Evaluative Conditioning

Evaluative Conditioning refers to the change in the valence of a stimulus resulting from pairing this stimulus with another one (De Houwer, 2007; De Houwer et al., 2001).⁹ In a paper often seen as the seminal paper of modern research on Evaluative Conditioning, Levey and Martin (1975) first asked participants to sort pictures according to whether they liked them. Then, they showed participants neutral stimuli along with the most liked and disliked stimuli. Each neutral stimulus was always paired with either a positive or a negative stimulus. The results of this experiment showed that stimuli paired with the most liked picture were evaluated more positively than the ones paired with negative pictures, which is coherent with the acquisition of new evaluative response due to the pairing. Since this seminal paper, Evaluative Conditioning has become a robust finding of the psychological literature (Hofmann et al., 2010).

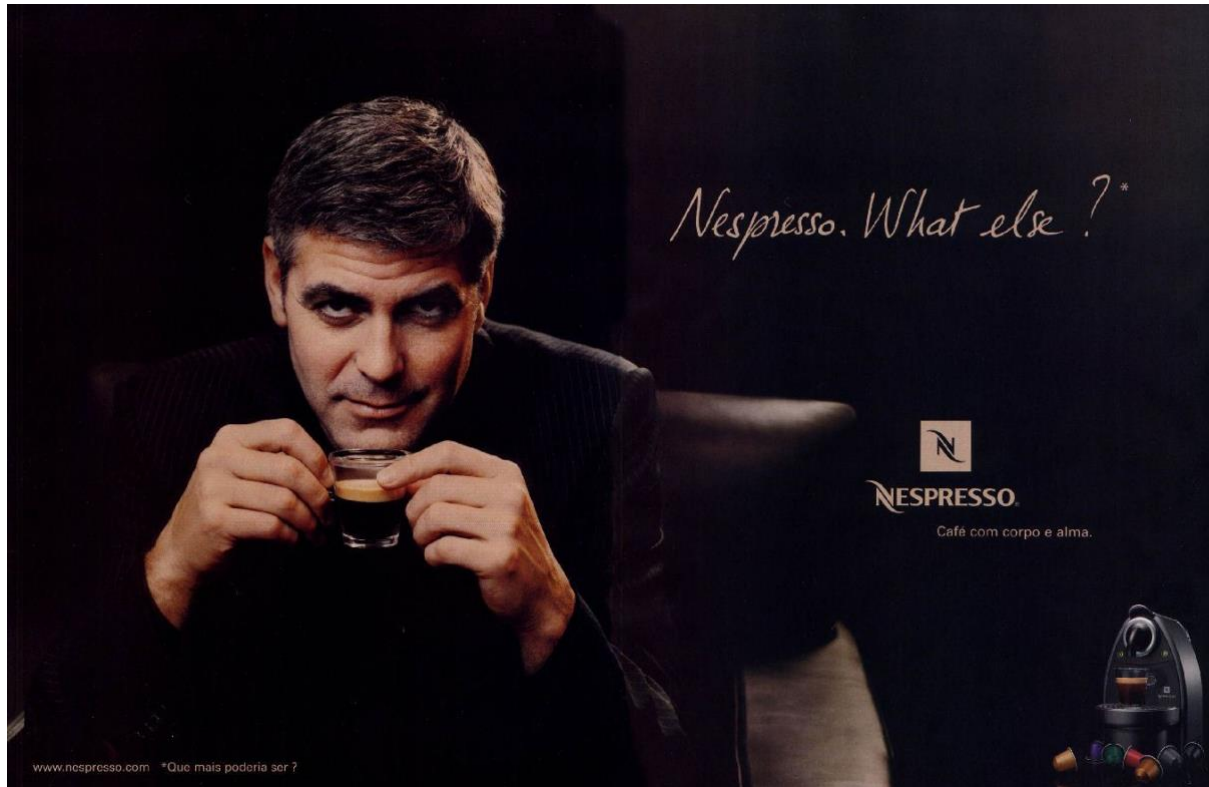
⁹ Note that Evaluative Conditioning was sometimes described as the mechanism which produces the associations in memory. At the time, it could be described as the mental association resulting from repeated pairing between stimuli. However, a functional definition is a better alternative as it does not narrow the scope of the research and contains less theoretical assumptions (De Houwer et al., 2013).

Since the first investigations of this phenomenon, Evaluative Conditioning has become one of the go-to paradigms used in implicit social cognition research. The core principles have stayed the same with the investigation of the effects resulting from the pairing of two stimuli. While the paradigm is one of the oldest in implicit social cognition, it is still relevant and still actively used to investigate theoretical questions in implicit social cognition research (e.g., Heycke & Gawronski, 2020; Hütter & Sweldens, 2018). The popularity of Evaluative Conditioning has largely crossed the boundaries of this research domain and one can find its application in domains such as marketing or clinical psychology.

Some marketing campaigns share a lot of similarity with Evaluative Conditioning paradigms. A recent example is the Nespresso advertising campaign where George Clooney appeared in every single ad Nespresso aired (see Figure 4.1; see Unkelbach & Högden, 2019). From the perspective of Evaluative Conditioning, assuming that George Clooney has a positive valence, repeatedly being exposed to an ad pairing George Clooney and a coffee capsule should result in the coffee capsule gaining George Clooney's positive valence. Reports suggesting pairing of stimuli as a viable paradigm to increase products like have been present for some time now (e.g., Gorn, 1982). Evaluative Conditioning even escaped advertisement campaigns to join political ones and the US Presidential Campaign opposing Al Gore to George W. Bush have brought a campaign clip where Democrats were associated with the word "rats" (Berke, 2000). Beyond these examples, which could be considered as controversial, Evaluative Conditioning has also been used with clinical applications.

Figure 4.1

Example of Advertisement Where Two Stimuli Are Paired



One of the problems related to substance addiction is that it sometimes comes with cognitive biases triggering specific attentional or evaluative response toward the addictive stimuli (R. W. Wiers et al., 2013). Because these biases may play a role in maintaining addictions, researchers have developed techniques to modify these cognitive biases (for a meta-analysis, see Boffo et al., 2019). Evaluative Conditioning is one of these tools which has been used as an add-on to therapies in the addiction context. For example, Houben et al. (2010) paired alcohol-related stimuli in an Evaluative Conditioning procedure which resulted in participant acquiring more negative evaluative responses toward alcohol.

Regarding the underlying mental construct assumed to operate in Evaluative Conditioning, the three models of evaluation that we describe above (i.e., the SEM, the APE model, and the single process propositional model; De Houwer, 2014, 2018; Gawronski & Bodenhausen, 2007, 2014a, 2018; McConnell & Rydell, 2014; Rydell & McConnell, 2006) assume that the repetition of the pairing will result in an association between concepts in memory. The main difference between the dual models and the single-process propositional models would be the nature of the association (mere association or proposition) and the operant condition necessary for the emergence of this association. As mentioned above, it is clear that the repeated experience principle is endorsed in Evaluative Conditioning, but it is not the only paradigm where the idea of a repeated exposure is implemented.

A reader familiar with the (classical) conditioning literature could be surprised to see how close the ideas implemented in this procedure is similar to the pioneering work on conditioning. Research on conditioning inspired several implicit social cognition paradigms. A notable specificity of Evaluative Conditioning paradigm is the lack of active behavior from the participants. Usually, participants taking part in Evaluative Conditioning are merely presented different kinds of stimuli, some which have a positive or negative valence, and some of which researchers are trying to change the evaluation. This might seem similar to the way Pavlov conditioned his dogs to react to the sound of a bell with food (in fact, one will find the first references to the Evaluative Conditioning paradigm under the name of classical conditioning of evaluative response; Levey & Martin, 1975). While Evaluative Conditioning finds its roots in classical conditioning, operant conditioning has inspired other paradigms. Operant conditioning differs from classical conditioning because it requires participants to perform an action toward a stimulus rather than being merely exposed to it. The next paradigm we will review—and the main paradigm of interest in the present dissertation—is

one of these paradigms. In the next section, we will introduce approach and avoidance training, an experimental procedure used in implicit social cognition research in which participant must repeatedly approach and avoid stimuli.

Approach and Avoidance Training

Evaluative conditioning is not the only paradigm which adopts the repeated experience of an event in its design. In approach-avoidance training, rather than passively perceiving co-occurrences between stimuli, participants have to repeatedly approach and avoid stimuli. In the end, the action that participants must perform and the stimulus that is presented both co-occur. In the present dissertation, we decided to focus on this paradigm. We will first review the emergence of this paradigm before discussing how it fits with the evaluation acquisition model we review above (i.e., the SEM, the APE model, and the single process propositional model).

Approach and avoidance behaviors are one of the most important behaviors of living organisms (Strack & Deutsch, 2004). From a survival perspective, it is critical to avoid what is bad for them and to approach what is good for them. An organism unable to do so would be unlikely to survive very long. Research in psychology investigating the approach-avoidance behaviors wonders whether property of the environment could trigger approach and avoidance in human beings and, consistent with the central aspect of approach and avoidance in our survival, Chen and Bargh (1999) show that stimuli with positive or negative valence could automatically trigger the approach and avoidance behaviors (in this experiment, the behaviors appeared unintentionally; for the seminal paper, see Solarz, 1960). Since this first paper, a meta-analysis now suggests that people are indeed quicker when they have to perform an approach movement toward something positive and an avoidance movement toward something negative compared to the opposite (Phaf et al., 2014). Overall, the approach-avoidance literature suggests that some aspects of the environment ease

approach and avoidance movements, but, interestingly for the present dissertation, it appears that performing approach and avoidance movements also changes the way we evaluate our environment.

In one of the first investigations of how movements could impact evaluation, Cacioppo et al. (1993) asked their participants to evaluate whether Chinese ideograms were pleasant while performing a specific motor action. In this experiment, participants had either to perform an arm flexion—as if they were bringing an object toward themselves—during the evaluation of half of the stimuli or an arm extension—as if they were pushing an object away from them—during the evaluation of the other half of the stimuli. As it could be expected if we assume a link between approach and positive things and avoidance and negative things, the Chinese ideograms were evaluated as more pleasant when participants performed a flexion movement during the evaluation compared to when participants performed an extension movement. These results, which were replicated twice in Cacioppo et al.’s original paper (Cacioppo et al., 1993; but for some limitations, Centerbar & Clore, 2006) provided the initial evidence that movement could impact evaluation.

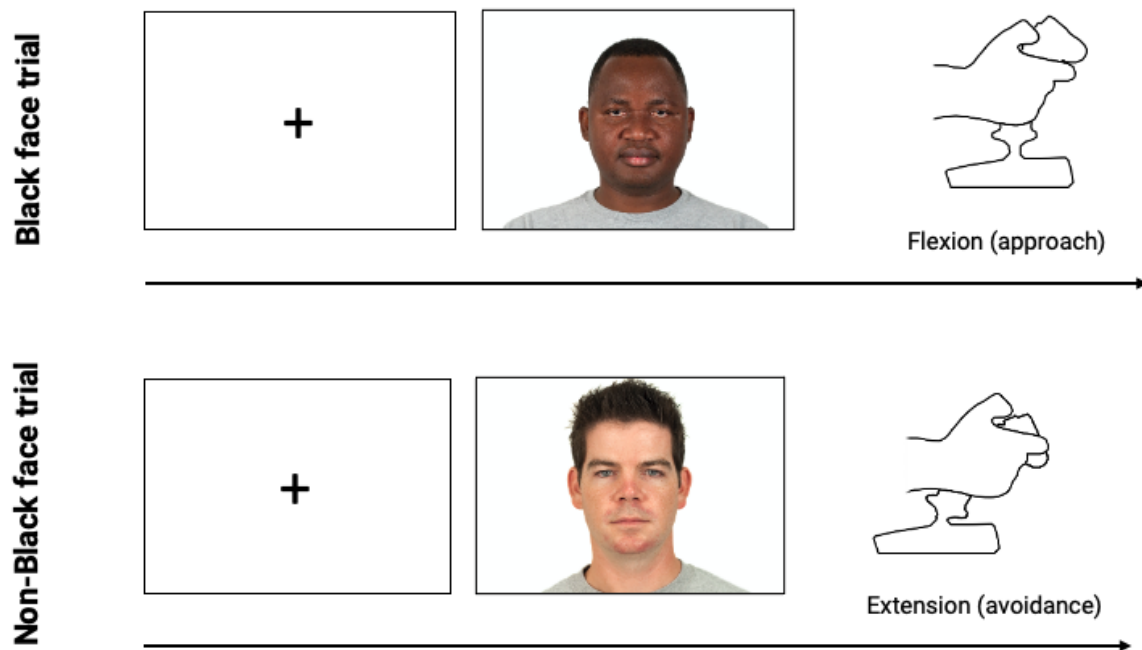
Here, evaluative response being direct, any of the evaluation acquisition model we presented (i.e., the SEM, the APE model, and the single process propositional model; De Houwer, 2014, 2018; Gawronski & Bodenhausen, 2007, 2014, 2018; McConnell & Rydell, 2014; Rydell & McConnell, 2006) could explain these results by making the assumption that performing the movement resulted in the formation of propositions impacting the evaluative response. However, because participants simultaneously perform the flexion and extension movement and the evaluation, it is not clear whether the movement impacts something related to the stimulus. In other words, a flexion movement could result in a proposition associating the mere movement to a positive evaluation rather than in a proposition

associating the movement to the stimulus resulting in a positive evaluation. If true, this explanation would reduce the relevance of this procedure, as it could not be used to change subsequent use of the stimuli. In other words, such paradigm would be helpful to change the evaluation of a stimulus at a given time, but not in the long run. Following work on approach and avoidance movements and their effect on evaluation ruled out this explanation: As we will see, approach and avoidance movements can be used to change how a stimulus is evaluated even when the movement is no longer performed.

Approach and avoidance training effects, at a functional level, are defined as the change in evaluative response to resulting from the performance of an approach or avoidance movement toward a stimulus. In an approach and avoidance training, contrary to the procedure used in Caccioppo et al.'s (1993) work, approach and avoidance movements are not performed during the evaluation, but before. In one of the most influential studies illustrating this kind of effect, Kawakami et al. (2007) successfully changed how non-Black participants evaluated Black people (on an IAT; see Figure 4.2). By repeatedly making participants approach photographs of Blacks, Kawakami et al. (2007) showed that movements influence subsequent indirect evaluations. In their original paper, Kawakami et al. (2007) showed that these movements impacted participants' IAT performance as well as their immediacy behaviors toward Black people such as body orientation toward them or seating (see Word et al., 1974). This intervention gave birth to the approach and avoidance training procedures, which popularity grew outside of intergroup relationships.

Figure 4.2*Time Course of Trials in a Black Face Approach Training*

Black face approach training



Note. By making participants repeatedly approach Black people's faces and avoid non-Black people's ones in a Black face approach training, Kawakami et al. (2007, Study 1) changed their participants' performance on an IAT. Participants were faster to perform the IAT when the key used to categorize positive words and Black faces was the same compared to a control condition where participants did not perform any approach and avoidance movement, a pattern usually associated with a more positive evaluation of Black people.

Beside intergroup relationship (Kawakami et al., 2007; Phills et al., 2011), approach and avoidance training has been used in other domains. Researchers applied these training to change women's math evaluation (Kawakami et al., 2008), to reduce alcohol use (A. M.

Hahn et al., 2019), to reduce cigarette use (Wittekind et al., 2015), to reduce cannabis use (Jacobus et al., 2018; B. J. Sherman et al., 2018) to encourage people to adopt healthier eating behavior (Ferentzi et al., 2018; Van Dessel et al., 2018b), to promote safe sexual behavior (A. M. Hahn et al., 2019), to reduce the tendency to gamble (Boffo et al., 2017), or even to reduce depression symptoms (Becker et al., 2019). The idea underlying these interventions is simple: Participants must perform an action linked to approach or avoidance in front of a set of stimuli and, typically, researchers make them approach things considered positive (e.g., math) and avoid things considered negative (e.g., cigarettes) to change the participants' behavior to a more adaptive one (e.g., persist during math exercise or quit smoking). As we will see, the theoretical explanation regarding the effectiveness of these interventions is quite similar to the one explaining Evaluative Conditioning effects. Note that we will focus on their effect on indirect evaluations rather than direct ones.

Approach-Avoidance Training and Theoretical Models of Evaluation

Along with Evaluative Conditioning paradigms, approach and avoidance training paradigms belong to the category of paradigms where participants repeatedly experience something (i.e., a pairing between a conditioned stimulus and an unconditioned stimulus; a pairing between a movement and a stimulus) and which result in an indirect evaluative response change (e.g., IAT performance with an approach-avoidance training; Kawakami et al., 2007). Dual process theories explain these effects by endorsing the Hebb's rule of "what fire together, wire together" in their operating principles (see Gawronski & Bodenhausen, 2006; McConnell & Rydell, 2014).¹⁰ In these models, the repeated co-occurrences between

¹⁰ Note that, because the APE model assumes the existence of indirect learning routes, one can learn an indirect evaluative response without the repeated co-occurrence between a movement and a stimulus in approach and avoidance training (see The APE Model section).

the movement and a stimulus result in an associative link between the movement and the stimulus. This mere association in memory will later cause the activation of the associated movement when the stimulus is perceived. Critically for the relevance of approach and avoidance training, one can expect that the reactivation of an approach or avoidance movement is not neutral in valence. One is expected to approach positive things more often than negative ones, the opposite being true for avoidance (Strack & Deutsch, 2004). Evidence in the literature supports the hypothesis that performing an action with a positive or negative outcome toward a stimulus results in a change of evaluation of the stimulus (Eder et al., 2019).

Regarding the operating principles they endorse, single-process propositional models account for the effects of approach and avoidance training differently. In a recent review, Van Dessel et al. (2018a) introduced a single-process propositional model for evaluation acquisition following approach and avoidance training. Rather than assuming that new indirect evaluations emerge as a result of a mere co-occurrence between a movement and a stimulus, Van Dessel et al. (2018a) suggest that participants need to draw an inference¹¹ a) between the stimulus and the movement (e.g., “I am approaching person A”) and b) between the approach movement and an evaluation (e.g., “I usually approach what I like”) for a new evaluation to emerge. Once these two propositions are acquired, it is likely that c) participants draw an inference between the approached stimulus and an evaluation (e.g., “This person A must be a good person”). One of the reasons why single-process propositional models stand

¹¹ Van Dessel et al. (2018a) use the term inference to describe a proposition emerging as a result of another proposition. In this section, we will use the two terms interchangeably. For a strict definition of what a proposition is, see the Propositional Processes section.

out from dual models is in how they account for instruction-based paradigms, as we will review later. Another one is the operating conditions each type of model adopts.

Accounting for indirect evaluative acquisition with associative processes involves endorsing operating principles, but also operating conditions. Approach and avoidance training effects were notably observed in suboptimal condition (i.e., with a subliminal presentation of the stimuli to approach and avoid; C. R. Jones et al., 2013; Kawakami et al., 2007), suggesting that learning can occur without any intention of learning and without people being aware that a learning is occurring. However, the idea that indirect evaluation acquisition through approach and avoidance training relied on automatic processes was recently challenged. These discussions emerged following new research (e.g., Van Dessel et al., 2015), but also failed replication of studies suggesting that one could acquire an evaluative response in an approach and avoidance training when the stimulus is presented in a suboptimal fashion (Van Dessel, De Houwer, Roets, et al., 2016).¹² Critically, the operating conditions implied by associative processes and propositional processes differ. Indeed, while a model à la McConnell and Rydell (2014) would argue that one can acquire an evaluative response without being aware of an association between a movement and a stimulus, the inferential account would not make this prediction (Van Dessel et al., 2018a; see also De Houwer, 2018).

To sum up, regarding the contribution of repeated experience in indirect evaluative learning through approach and avoidance training, every model suggests a specific contribution. As discussed in the section dedicated to the SEM and the APE model, the two

¹² Van Dessel et al. (2016) do not merely fail to replicate past studies. They provide strong evidence for a null hypothesis which makes their replication studies more informative than mere non-significant results (Lakens et al., 2018).

models suggest that repeated experience is an important principle, as it is the main ingredient of associative processes—the process dedicated to indirect evaluative responses (Gawronski & Bodenhausen, 2007; McConnell & Rydell, 2014; Rydell & McConnell, 2006). While the SEM assumes that indirect evaluative learning through associative processes is not possible without repetition, the APE model assumes the existence of an indirect route in which propositional processes impact the associative network. Although the APE model assumes that we can learn an indirect evaluative response through propositional processes (Gawronski & Bodenhausen, 2006, 2007, 2014a), repeated experience is still supposed to influence indirect evaluative response acquisition, as propositions are more likely to emerge with repetitions (Van Dessel et al., 2018a). In the end, every model that we review (be it dual-process or single-process) assumes a specific role for repeated experience in evaluative learning. Despite this assumption, some experimental paradigms used in implicit social cognition research do not include this ingredient.

Learning Through Instructions

Every paradigm reviewed so far has in common participants being repeatedly exposed to an experience. Within the context of Evaluative Conditioning, it is a co-occurrence between a conditioned and an unconditioned stimulus, within the context of approach and avoidance training, a co-occurrence between a stimulus and the movement. While the conditions of exposure differ between the different methods, some of the paradigms adopted by the implicit social cognition literature embrace the idea of repeated experience as an active ingredient. This observation is coherent with the roots of implicit social cognition suggesting that we could capture something which was progressively encoded over time (Fazio et al., 1986; Greenwald & Banaji, 1995). As we will see, however, recent work stands out from the studies adopting repeated-experience-based paradigms.

In an experiment, De Houwer (2006b, Experiment 1) told his participants that they would take part in a study where they would see several non-words (e.g., BAYRAM) as well as images (e.g., flowers). They were informed that these stimuli would be presented sequentially and that the non-words would be diagnostic of the valence of the following image (e.g., if you see BAYRAM, a positive image will appear; if you see UBIDON, a negative image will appear). Critically, participants never actually performed this task, but took an IAT instead. In its design, this study looks a lot like an Evaluative Conditioning experiment, but without the active ingredient of the paradigm—the repeated exposure. Right after receiving the mere Evaluative Conditioning instruction, participants passed an IAT where they categorized the non-words with the same response keys used to categorize positive and negative words. The results showed that it was easier for participants to perform the task when they used the same key to categorize the non-words diagnostic of positive valence image and positive words than the opposite. While this experiment was designed to test whether conscious propositional knowledge could influence an IAT,¹³ it is an illustration of how mere instructions can impact indirect evaluative responses.

Since this experiment, other experiments illustrated the effects of instruction-based procedures. In an experiment, Gast and De Houwer (2012) exposed their participants to a set of positive and negative stimuli—some images from the International Affective Picture System (IAPS; Lang & Bradley, 2007). These positive and negative stimuli were not presented alone, however, but with either one of two rectangles. These rectangles could be labeled with the number “1” or the number “2” and were always paired with stimuli with a same valence. For example, a participant could be exposed to positive stimuli accompanied

¹³ At the time, the idea that the IAT captured something outside of individuals’ consciousness was prevalent in the literature.

by the 1-labeled rectangle and negative stimuli accompanied by the 2-labeled rectangle. The critical instructions appeared right after the pairing phase. Participants were told that each of the 1- and 2-labeled rectangles was actually hiding an image. While participants did not experience any pairing between the positive and negative stimuli and the images hidden by the 1- and 2-labeled rectangle, they were led to believe that a pairing happened. Critically regarding the role of instructions in evaluation acquisition, an IAT revealed that participants developed specific evaluative response to stimuli they thought had been paired with positive and negative pictures—even if they did not experience the pairing. These results illustrate that one can acquire an indirect evaluative response in an Evaluative Conditioning lookalike procedure where participants do not actually experience the pairing between the stimuli.

Interestingly, instruction effects generalize to approach-avoidance lookalike procedures. In their work, Van Dessel et al. (2015) investigated whether merely receiving approach and avoidance instructions would make participants acquire new evaluative responses. In their experiment, they showed participants either novel social groups (Niffites and Luupites; Experiment 1; see Gregg et al., 2006) or groups with which participants were familiar (Black and White people; Experiment 2; see Kawakami et al., 2007). Participants received the instructions to approach one group and avoid the other (or the opposite) and were then administered an IAT where they had to categorize positive words and negative words along with the stimuli that belonged to the category they were told to approach and avoid. Participants' IAT performances revealed the acquisition of new indirect evaluation when they received instructions to approach and avoid novel stimuli (Experiment 1); A significant effect of instructions was not found when participants were familiar with the social groups (Experiment 2; see also, Van Dessel et al., 2020). These findings, which were replicated in other studies, suggest that instructions can easily form new indirect evaluations.

At first glance, these results might be surprising. Historically, indirect measures were developed to make up for some of the limits of direct evaluations. With these indirect evaluative measures (e.g., Fazio et al., 1986; Greenwald et al., 1998), researchers wanted to capture an evaluative response emerging automatically. By doing so, they were hoping to overcome limits of direct evaluative responses such as the ability for a participant to provide an answer with a motivation other than accuracy. In developing these measures, most of the theoretical models adopted a view where indirect evaluations captured the results of repeated co-occurrence happening over time and slowly learned. Paradigms based on instructions do not really fit this description of a paradigm based on Hebb's rule. Still, when we take a look at what actually are the predictions of theoretical models of evaluation such as the APE model or a single process propositional model¹⁴, there is room for indirect evaluative response.

If both instructions and repeated experience can make someone acquire an indirect evaluative response, however, one critical question remains. What is the actual contribution of repeated exposure when it comes to indirect evaluation? This is the questions we want to address in the present dissertation.

It is important to remember that the question of the contribution of repeated experience emerged because of recent advances in implicit cognition (e.g., Gast & De Houwer, 2012; Van Dessel et al., 2015). In 2016, at the time we started this work, to the best of our knowledge, no other line of research investigated the role of repeated experience by comparing repeated-experience- and instructions-based paradigms (but for some work

¹⁴ It has to be noted that the single-process propositional models were partly developed with the goal to account for instructions effect (De Houwer, 2009). Therefore, it should not be a surprise that they do indeed account for the instruction effects.

comparing the effects of mere association to the effects of verbal statements, see Peters & Gawronski, 2011; Zanon et al., 2014). Still, since 2016, several teams of researchers decided to compare these two procedures and, nowadays, there are several papers tackling this question in the literature (e.g., Kurdi & Banaji, 2017; Rougier et al., 2021; Van Dessel et al., 2020). Because these papers are not critical to understand the challenges of this dissertation, we decided not to review them in the introduction section of this manuscript. Instead, to offer a comprehensive view on this question, we discuss these lines of research and how they fit with the literature in the general discussion.

Chapter 5

Overview of the Present Dissertation

In the introduction section of the present manuscript, we reviewed some of the critical ideas embedded in the implicit cognition literature. In doing so, we showed that one of the core assumptions of implicit social cognition theories is that the repetition of an experience contributes to the development of indirect evaluative responses. Recent advances, however, showed that merely providing participants with the instructions of repetition-based paradigm is sufficient for an evaluative response to emerge. While most of the modern theories of evaluation can account for these instruction effects, one could still wonder whether the repetition-based paradigms outperform the instructions-based paradigms. As discussed in the introduction, each of the theoretical models that we reviewed predicts at least some situations where they should. In the present dissertation, we argue that theoretical models of evaluation predict a superiority of repeated-experience-based procedures over instruction-based procedures regarding their impact on the acquisition of indirect evaluative response.

Our response to the question of the difference between repetition-based procedures and instructions-based procedure was experimental. Because approach-avoidance behaviors represent one of the most primitive functions of an organism (Strack & Deutsch, 2004), we decided to focus on this paradigm. We focused on approach-avoidance procedures. In the experimental part of this manuscript, we will report each of the experiments that we conducted. An overview of each of these experiments can be found at the end of this section (see Table 5.1, p. 92). These experiments will be organized in five different chapters.

In the first empirical chapter, **Chapter 6**, we developed an approach avoidance paradigm based on the VAAST procedure (Rougier et al., 2018). In this section, we introduce

the repetition-based paradigm that we will use in the present dissertation. We benchmarked this paradigm by comparing it with a procedure used in the approach-avoidance paradigm literature, namely, the Joystick task (Experiment 1a). We also assessed the usability of such paradigm on online samples (Experiment 1b), and, finally, offer a preregistered replication of an experiment describing a math-approach training designed for women with low identification to math (Experiment 2; Kawakami et al., 2008).

We will report our first assessment of the difference between repetition-based and instructions-based paradigms in **Chapter 7**. We conducted two experiments which included our VAAST-based approach-avoidance training and an approach-avoidance instructions procedure. In Experiment 3, we tested whether these procedures can form evaluative responses on an Affective Misattribution Procedure (Payne et al., 2005) and whether the magnitude of the evaluative responses differed between the two procedures. Our objective was similar in Experiment 4, but this time, we adopted a within-subject design—our participants evaluated both stimuli that they had approached and had avoided and stimuli for which they merely had received instructions.

In **Chapter 8**, we will report our investigation of the effect of approach-avoidance training in an experiment adopting an incidental design. In a nutshell, in Experiment 5, we investigated whether approaching and avoiding words during a training resulted in the acquisition of an evaluative response, even if the said words were not mentioned in the instructions.

In **Chapter 9**, we investigated the effects of both approach-avoidance training and approach-avoidance instructions after a delay. In Experiment 6, participants engaged with either one of the two procedures and then took an IAT. They also took the same IAT a day after the approach-avoidance procedure. This allowed us to investigate whether the evaluative responses they had acquired as a result of the approach-avoidance training lasted longer than

the evaluative responses they had acquired as a result of approach-avoidance instructions. In Experiment 7, we replicated some of the findings of Experiment 6.

Finally, in **Chapter 10**, we examined people's lay belief about the two procedures. Given that approach-avoidance training is sometimes used within an application context, we considered investigating participants' outcome expectation worthwhile. In Experiments 8 and 9, we described the procedures to the participants and asked them to make predictions regarding the outcome of such procedure on an indirect evaluative response. Experiment 9, compared to Experiment 8, included an approach-avoidance training demo to help participants picture what it is like to be someone in a study investigating approach and avoidance procedures.

Table 5.1*Overview of the Experiments*

		Measure	SS.	Sample	OSF repository
Exp. 1a	VAAST approach-avoidance training benchmarked against the Joystick task	IAT	140	Lab	https://osf.io/vtbdh/
Exp. 1b	Online version of the VAAST approach-avoidance training	IAT	80	Online	https://osf.io/jv5ns/
Exp. 2	Replication of Kawakami et al. (2008)	IAT	203	Lab	https://osf.io/amq99/
Exp.3	Approach-avoidance procedures comparison	AMP	63	Lab	https://osf.io/nu9fg/ *
Exp.4	Approach-avoidance procedures comparison	AMP	104	Lab	https://osf.io/qty6w/
Exp.5	Incidental vs. Instrumental approach-avoidance procedures comparison	AMP	238	Lab	https://osf.io/5tcn4/
Exp.6	Comparison of the effects of approach-avoidance effects at D+1	IAT	408	Online	https://osf.io/etsmv/
Exp.7	Replication of Experiment 8 (D-Day only)	IAT	996	Online	https://osf.io/t4w3e/
Exp.8	Naive predictions of the difference between approach-avoidance procedures	Guessing	150	Online	https://osf.io/he2qu/
Exp.9	Naive predictions of the difference between approach-avoidance procedures (with demo)	Guessing	151	Online	https://osf.io/cf27h/

Note. This experiment was not preregistered. AMP = Affect Misattribution Procedure,

IAT = Implicit Association Test, SS = Sample size.

* Experiment 5 was a pilot experiment. It was included at the end of another experiment from our lab and can be considered underpowered.

PART II: EXPERIMENTAL PART

Chapter 6

A VAAST-based approach-avoidance training

The goal of the present chapter is twofold. We want to 1) validate an approach and avoidance training procedure based on the Visual Approach Avoidance by the Self Task (the VAAST; Rougier et al., 2018) and 2) conduct a replication of an experiment illustrating the effects of a math approach training procedure to increase score on a math identification IAT (Kawakami et al., 2008). The first goal will help us regarding the overarching goal of this dissertation, that is, investigating the effect of repeated experience when it comes to indirect evaluation acquisition.

As discussed in the introduction, we characterize approach-avoidance trainings as procedures in which individuals repeatedly approach and avoid stimuli. The VAAST is a procedure implementing approach and avoidance movements which was developed as an approach-avoidance tendency measure (Aubé et al., 2019; Rougier et al., 2018, 2020). In a nutshell, this task allows researchers to assess whether a stimulus is easier to approach than it is to avoid. To capture approach and avoidance tendencies, participants perform approach and avoidance movements of the whole self in a computerized environment. While this procedure is to be a robust measure of approach and avoidance, it was never used as an approach and avoidance training. We will use the present section as an opportunity to put the VAAST paradigm's approach-avoidance training potential to the test.

To assess the validity of the VAAST as an approach-avoidance training, we decided to replicate work conducted by Kawakami et al. (2008) in which participants' evaluative responses toward math as the result of a math approach training. Despite being relevant to the issue of the underrepresentation of women in STEM fields (Nosek et al., 2002, 2009),

Kawakami et al.'s (2008) experiment was never replicated. By doing so, we would 1) offer evidence that the VAAST can be used as an approach-avoidance training procedure and 2) update our knowledge about the effects of math approach training. Two experiments will precede the replication of Kawakami et al. (2008). Because the VAAST had never been used as an approach-avoidance training procedure before, a non-replication of Kawakami et al.'s (2008) could have been interpreted either 1) as evidence against the effectiveness of math-approach training or 2) as evidence against the effectiveness of the VAAST as an approach-avoidance training. The two experiments preceding the replication of Kawakami et al. (2008) will allow us to overcome the aforementioned limitation.

In this section, we will first discuss why replicating Kawakami et al.'s (2008) is an important contribution to the social psychology literature. Then, we will report the results from Experiments 1a & 1b which introduce the VAAST-based approach-avoidance training, and we will follow with the replication of Kawakami et al. (2008). We will report a meta-analysis on the effect of math approach trainings along with results of the replication. Note that this chapter is an edited version of a published manuscript (Batailler et al., 2021). Compared to the original manuscript, in this chapter, we include Experiment 1b which tests whether an online version of the VAAST can be used to form evaluative response. This online version of the paradigm could be useful to increase sample size and the ease of access of certain populations.

Implicit Identification and the Gender Gap in Science

Implicit identification is theorized as the determinant of spontaneous tendencies to associate oneself with concepts, people, or stimuli (Greenwald et al., 2002). Besides mediating the relationship between a stimulus and its indirect evaluation (see De Houwer et al., 2013) such as in the identification Implicit Association Test (identification IAT; Nosek et al., 2002), implicit identification mediates the relationship between a concept and some of its

related behavior. For example, implicit identification to psychology as a field predicts academic persistence in psychology (Roland et al., 2018). Therefore, in order to address problematic behaviors, researchers have been trying to target implicit identification in interventions (Kawakami et al., 2008; Phills et al., 2011).

One important area where social psychology researchers have been investigating the role of implicit identification is the gender gap in STEM fields (Wang & Degol, 2017). A large body of work has explored the role of identity-related process in the underrepresentation of women in STEM-related fields, showing, for example, that a salient female identity could impact women's performance in math (Spencer et al., 1999), even among lower elementary grades (Ambady et al., 2001). In a large-scale study, Nosek et al. (2009) went further by examining an indirect evaluation intended to measure implicit associations between male (female) and science (liberal arts) at the country level. Using data from the 2003 Trends in International Mathematics and Science Study (Gonzales et al., 2004), Nosek et al. were able to predict countries' gender gap in math achievement using scores from the IAT (Greenwald et al., 1998): The stronger the male-science association, the bigger the gender gap. The conclusion they drew out of this correlation was that the bias to associate males more with males than with females and the gender gap in math achievement would reinforce each other. That is, low representation of women in science would make it harder for women to identify with math and this dissociation would drive women away from math, ultimately making women underrepresented in this field. In such a context, making identification between women and math higher would create a positive feedback loop reducing the gender gap in science. This idea makes any intervention reducing the dissociation between women and math extremely valuable. The math approach training Kawakami et al. (2008) developed targeted this identification to math.

Increasing Math Implicit Identification with an Approach Training

In an influential paper (more than 122 citations according to a Google Scholar query), Kawakami et al. (2008) presented two studies investigating whether merely approaching math-related pictures could change women's math identification. Recruiting women who reported not liking math, they demonstrated that repeatedly approaching (avoiding) math-related pictures and avoiding (approaching) arts-related pictures had an impact on identification and attitude IATs (Greenwald et al., 1998; Nosek et al., 2002). In these studies, participants approached and avoided stimuli by respectively pulling a joystick (as if they were bringing an object toward them) and pushing it (as if they were pushing away an object). Approaching math (instead of avoiding it) reinforced the participants' self-identification with math and made them evaluate math more positively. We think this paper offers an important leverage to reduce the gender gap in science.

Now, with the goal to build upon this theoretical foundation, we first need to assess its evidential value, that is the strength of evidence that has been collected for this effect. One classic way to do so when enough studies have been conducted is to rely on a formal meta-analysis (even though there have been discussions about the efficiency to correct for the publication bias, see Carter et al., 2019)—the tendency to publish mostly significant findings—and for degrees of freedom in data analysis when studies have not been preregistered; (Simmons et al., 2011). Because, to the best of our knowledge, there is no work other than Kawakami et al.'s (2008) testing the math approach training effect, a formal meta-analysis is not an option. Fortunately, the literature provides more and more tools to assess evidential value when only few studies are available.

Assessing the Evidential Value in Favor of the Math Approach Training

To begin with, and somewhat at odds with an all or nothing interpretation of p values (i.e., significant or not with regard to the .05 threshold; Giner-Sorolla, 2016), one can often

infer stronger evidential value when p values are farther away from .05 (that explains that, for instance, Benjamin et al. suggested a .005 threshold for stronger evidence, 2018). Along the same line, Simonsohn (2014) developed an index (the p -curve) that relies on the principle that when an effect exists in the population, smaller p values should be more frequent than p values close to .05 (i.e., a downward slope). Such a tool, however, is not suited for single papers as assessment of evidence (Simonsohn et al., 2014). Beyond p values, information that contributes to the assessment of evidential value includes effect sizes (i.e., whether most of the effect sizes are of the same direction and how far they are from the null effect), sample sizes (i.e., studies with more observations have less chance to reflect Type 1 errors; Lakens & Evers, 2014; Schönbrodt & Perugini, 2013), and the precision of an effect size (i.e., how narrow a confidence interval is, a narrower one coming with more confidence regarding the existence of an effect—providing obviously that 0 is not included in this confidence interval). It is worth mentioning that methodological aspects can also strengthen evidential value, for instance whether studies were preregistered (notably because it constrains degrees of freedom in how data were handled and because severe testing bring more evidence for a hypothesis; Mayo & Spanos, 2006). Obviously, Kawakami et al. (2008) could not use this way to increase evidential value insofar as it was not used in social psychology at the time. Finally, a tool that can be used to assess evidential value, even when only few studies are available is the Bayes factor. Such index indicates how much we should update our belief that a hypothesis is true (usually compared to a null hypothesis) after observing empirical data (i.e., one or several studies; Jeffreys, 1961; Wagenmakers, 2007). Bayes factors thus makes it possible to determine the level of evidence for a hypothesis in a study or a set of studies. Importantly, instead of simply looking at whether a replication study produces a significant p value or not, such Bayes factors quantify how much one learns from that replication.

In order to assess evidential value in favor of the hypothesis of a math approach avoidance training effect on implicit identification, one can therefore study either the aforementioned information or compute a Bayes Factor based on Kawakami et al.'s (2008) data. For instance, one can observe that the critical p values were .060 and .048 (respectively for Studies 1 and 2) and that the effect size estimations were rather imprecise (notably due to very small and small sample sizes respectively), with confidence intervals including or being close to a null effect size (see Figure 6.5, p. 125). Those two observations might suggest only weak evidence at this stage. In line with this conclusion, the Bayes factor we computed for these two critical tests reveals what is usually coined “moderate” evidence for this effect ($BF_{10} = 7.36$; Jeffreys, 1961).^{15, 16}

Increasing the Evidence for the Math Approach Training

Because we think Kawakami et al.'s (2008) paper offers an important leverage to reduce the gender gap in science regarding Nosek et al. (2009)'s results, but suffers from a

¹⁵ Level of evidence a Bayes factor can provide usually encompass “anecdotal”, “moderate”, “strong”, “very strong”, or “extreme evidence” for either a null hypothesis or a specific alternative (Jeffreys, 1961).

¹⁶ Note that our computation of the Bayes factor differs from the recommendation of Field, Hoekstra, Bringmann, and Van Ravenzwaaij (2019) regarding the assessment of evidential value for the purpose of selecting replication candidate. Indeed, we decided to use all the information available regarding the math approach training effect, resulting in the inclusion of Kawakami et al.'s (2008, Experiment 1) marginally significant math approach training effect (i.e., $p = .060$). As a consequence, evidence in favor of the math approach training might be overestimated. When this marginal effect is not included, the evidential value for the math approach training effect is only “anecdotal” ($BF_{10} = 1.70$; Jeffreys, 1961).

quite modest evidential value, conducting a close replication is useful as it could strengthen the level of evidence for this effect. A successful replication would also support the idea of a generalizable effect notably by using a different task, different stimuli, and a different population. In doing so, we will make sure to increase the sample size and to preregister the study, as this would increase the evidential value of the replication if the effect truly exists. In addition, we will adopt a different approach and avoidance training task (Rougier et al., 2018).

In their original experiments, Kawakami et al. (2018) adopted the Joystick task (without visual feedback) as an implementation of approach and avoidance. This task relies theoretically on the meaning associated with arm movements (for a seminal work in this area, see Cacioppo et al., 1993; Chen & Bargh, 1999): Arm flexion being associated with approach (bringing something toward themselves) and arm extension being associated with avoidance (pushing something away). Other studies, however, have shown that arm flexion could be associated with avoidance (e.g., withdrawing one's hand from something) and arm extension with approach (e.g., approaching one's hand from something; Paladino & Castelli, 2008; Seibt et al., 2008). This suggests that those mappings are not hardwired at the motor level, but are open to cognitive interpretation (Seibt et al., 2008). Consistent with this idea, at least when studying approach/avoidance as a measure (to demonstrate that people are, for instance, faster to approach positive stimuli and avoid negative stimuli than to perform the reverse actions), effects seem more robust when approach/avoidance operationalizations rely on movements of the whole self (where moving closer and away almost always means approaching and avoiding respectively) instead of arm movements (Krieglmeyer & Deutsch, 2010). This difference with movement of the whole self is even larger for arm movements operationalizations which do not use visual feedback in addition to arm movements (Krieglmeyer & Deutsch, 2010).

For these reasons, even though the current work studies approach/avoidance as a training and not as a measure, we decided to rely on an approach/avoidance operationalization that simulates visually movement of the whole self: the Visual Approach/Avoidance by the Self Task (VAAST; Rougier et al., 2018). In this task, participants receive instructions to approach and avoid categories of stimuli (e.g., “approach positive words and avoid negative ones”) appearing in a virtual environment (e.g., a street), to do so, they have to press either an “approach” or “avoid” response key. Pressing the “approach” response key triggers an animation on the screen corresponding to an approach visual flow: The point of view within the environment gets further and the stimulus in front of the point of view gets bigger; pressing the “avoidance” response key triggers an avoidance visual flow.

By adopting the VAAST rather than the Joystick task for an approach/avoidance training, we adopt a task that emphasizes visual feedback in approach and avoidance movements rather than motor activation. In both procedures, participants receive the instructions to approach and to avoid different classes of stimuli. However, the Joystick task relies (theoretically) on the movement the participants perform while the VAAST focuses on the approach/avoidance visual feedback participants’ responses entail. Because the latter are unambiguously mapped to either an approach or avoidance movement (Rougier et al., 2018), we think it is a safer implementation of approach and avoidance for an approach and avoidance training.¹⁷ Another possible added value could be that, even though it will not be

¹⁷ By safer, we mean that it is less likely with the VAAST implementation that the ambiguity of approach and avoidance implementation interferes with the training. Note that we consider this interference nonsystematic and more likely to occur only in specific condition (e.g., when

the case in the current studies conducted in the lab, the VAAST does not necessitate any particular material (like a joystick) and can easily be applied to online experiments (Aubé et al., 2019). Using the VAAST in order to replicate Kawakami et al.'s (2008) study, however, is viable only under the assumption that the VAAST can produce results at least as large (as a training) as the Joystick task. We test this assumption in Experiment 1a.

Overview of the Experiments

In this manuscript, we conducted a preregistered replication of Kawakami et al.'s (2008) by 1) substantially increasing the sample size and 2) using the VAAST as an implementation of approach and avoidance (Rougier et al., 2018). We focused on the approach and avoidance training effect on identification IAT because Roland et al. (2018) showed that this measure is a good predictor of academic persistence. In a pilot experiment, we first tested whether the VAAST could be used to form indirect evaluations (i.e., on an identification IAT) and we checked whether it performed at least as well as the Joystick task procedure, as implemented in Kawakami et al. (2008). This comparison was relevant, because in the event that we did not replicate Kawakami et al.'s (2008) results, doubts could be raised that it was due to our use of a different training task. Then, we replicated Kawakami et al.'s (2008) experiment with the VAAST, comparing a math approach training and a math avoidance training among women with low identification to math (see the replication recipe in Appendix 2; Brandt et al., 2014). Finally, we conducted a meta-analysis including our

it is easier for the participant to perform a specific movement compared to another for a target category; for similar reasoning, see Krieglmeyer & Deutch, 2010). That explains why we did not formulate the hypothesis that the VAAST should necessarily produce a larger effect than the Joystick task in the Pilot Study (see our pre-registration).

replication as well as Kawakami et al.'s (2008) results. We report all measures, manipulations, and exclusions.

Experiment 1a

The goal of the Experiment 1a was twofold. First, we wanted to test whether a VAAST-based training could form a specific evaluation (on an identification IAT). Second, we wanted to make sure that, if any difference between the VAAST-based training and the Joystick task training used by Kawakami et al. (2008), it would be negligible. If, for any reason, the Joystick task training effect was larger in this pilot experiment, it would have raised concerns regarding our goal to replicate Kawakami et al.'s (2008) work using the VAAST. We report all measures, manipulations, and exclusions. Material, data, and the analysis script, and the preregistration can be found online at: <https://osf.io/vtbdh/>.

Method

Participants, Design, and Sensitivity Analysis

One hundred and fifty ($M_{age} = 20.10$, $SD_{age} = 2.48$, 127 women) psychology students from a French university took part in this study in exchange for course credit. We relied on a 2 (training: Luupite approach vs. Niffite approach) by 2 (training task: Joystick task vs. VAAST) between-participant design.

The published literature suggests that the expected effect of approach and avoidance training is large ($d_s = 1.04$; Van Dessel, Gawronski, et al., 2017). With the sample size of 140 participants we preregistered, we had enough power to detect the training effect (i.e., $1 - \beta > .99$). Moreover, we had enough power to conduct a negligibility test for the difference between the Joystick task and the VAAST (i.e., $1 - \beta = .80$).

Remember that we do not expect large differences in the effectiveness of the Joystick task and the VAAST. Negligibility tests allow considering an effect as negligible if the estimated effect size is smaller than an a priori limit. In other words, to ensure that the

difference of effectiveness between the Joystick task and the VAAST was not large, we set an effect size interval in which the estimated effect of our study was likely to fall, if no effect existed in the population (e.g., 80% of the time). If the estimated effect in our study was outside this interval, we could not consider the difference between the Joystick and the VAAST negligible, even if the difference between both was not significant (these results would be considered as non-informative). This analytic strategy is used to address pitfalls regarding interpretation of non-significant effect as null. In our preregistration, in order to satisfy power requirements suggested by Cohen, we used an effect size of $\eta_p^2 = .013$ as our threshold (for an introduction on the concept of smallest effect size of interest, or SESOI, see Lakens, 2017).

Procedure and Material

When arriving at the lab, participants were led to an individual cubicle. The experimenter then explained that the experiment would be about movement and response time and that they would engage in multiple tasks for a total time of 25 minutes. All the instructions were displayed on the computer screen. The experimenter also informed participants that they might have to use experimental material during the experiment and that this material would be located behind the computer screen. The experimenter was blind to participants' condition. Participants were seated at approximately 70 cm from a 24-inch screen. The experiment was programmed using E-prime.

Approach and Avoidance Instructions. Two lists of words (i.e., the Luupite and the Niffite list; see Gregg et al., 2006) were presented to the participants. They were instructed that every word in the Luupite list ends with -lup and had two consecutive vowels (e.g., maasolup, tuuralup) and every word in the Niffite list ends with -nif and had two consecutive consonants (e.g., cellanif, otrannif). An exhaustive list of the stimuli can be found in Appendix 3. Participants' task for the first part of the experiment was to approach the words

of the one list and to avoid the words of the other. Half of the participants was randomly assigned to approach words from the Luupite list and avoid the one from the Niffite list, the other half had to do the opposite.

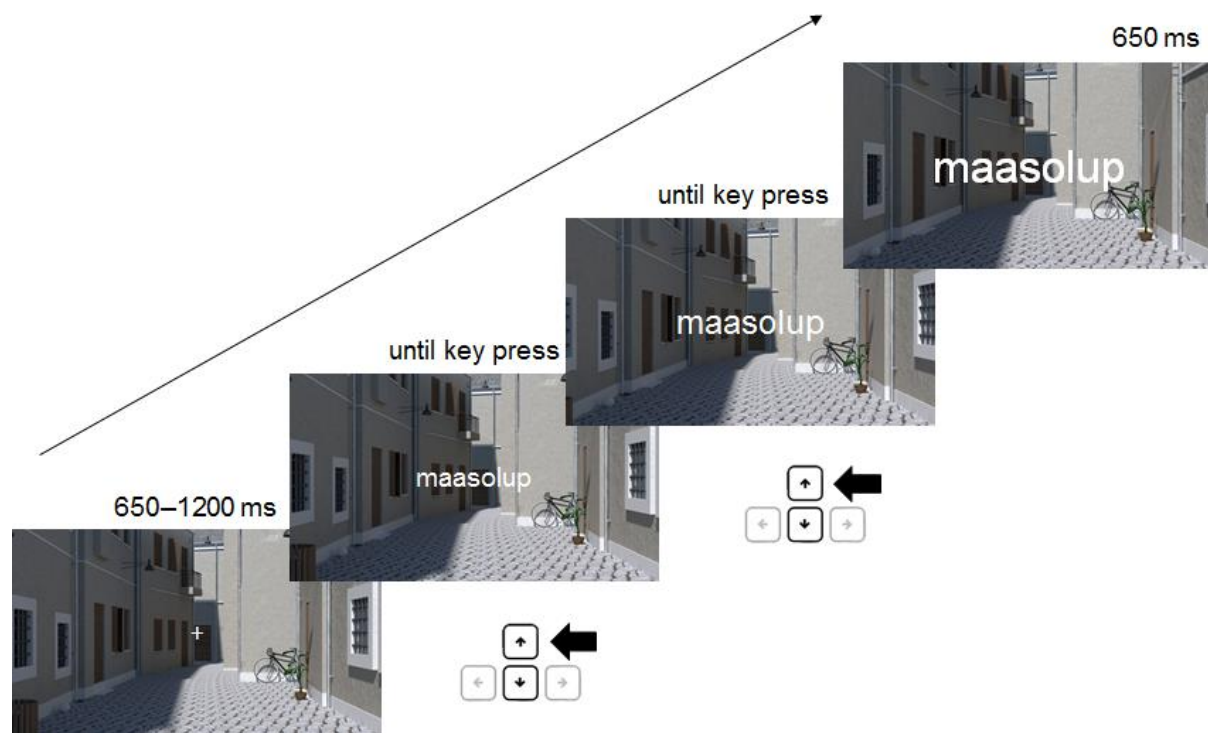
Approach and Avoidance Training. After receiving the instructions, participants performed the approach and avoidance training. They were randomly assigned either to the Joystick task training or the VAAST training.

Joystick Task Training. Participants performed an approach/avoidance task identical to the one used by Kawakami et al. (2008) but for the stimuli used. To approach the target word, participants pulled the joystick toward themselves (as if they were grabbing the word toward themselves) and to avoid the target word, they pushed the joystick away from themselves (as if they were pushing the word away).

For each trial, participants saw a fixation cross for a random duration (between 650 ms and 1200 ms) followed by a target word at the center of the screen. When the target word appeared, participants either pulled the joystick toward themselves or pushed away the joystick from themselves. In case of incorrect response, the “Error!” message appeared in a gray box and participants then had to start the trial over. Participants started by a practice block which consisted of 20 trials and then, completed the main block of 200 trials; one half of the trials was related to approach and the other one was related to avoidance. The training phase lasted approximately 10 minutes.

VAAST Training. Participants in the VAAST training condition were informed that they would have to perform approach and avoidance movements within a virtual environment on the computer and that they would have to use the keyboard arrow keys of the directional keypad to do so (see Figure 6.1). Participants pressed the up arrow when they had to approach the target word and pressed the down arrow when they had to avoid it.

Each trial started with a fixation cross for a random duration (between 650 ms and 1200 ms) followed by a target word at the center of the screen. When the word was presented, participants pressed the key corresponding to the movement they had to perform. When participants pressed the correct key, the screen simulated an approach/avoidance movement: The stimulus size, as well as the point of view within the environment, simulated a movement (see Rougier et al., 2018). Participants had to press the correct response key twice to complete a trial, each key press triggering an animation, as if they were walking two steps forward/backward within the environment. If participants made an error, as in the Joystick task condition, a gray alert box appeared on the screen informing them so and they had to start the trial over again. Participants started by a 20-trial practice block and then moved to a 200-trial main block; one half of the trials was related to approach and the other one was related to avoidance. The approach/avoidance phase lasted approximately 10 minutes.

Figure 6.1*Time Course of a VAAST Trial in the Experiment 1a*

Note. Participant has to approach Luupite words. Black arrows represent a key press.

Identification IAT. Immediately after the approach and avoidance training phase, participants completed an identification IAT (see Nosek et al., 2002). In this task, participants categorized words from the “Luupite” and “Niffite” lists as well as words related to the self (e.g., I, mine) and others (e.g., they, theirs). An exhaustive list of the stimuli can be found in Appendix 3. To complete this task, participants used a *Chronos* response box.

This IAT was identical to the one adopted by Kawakami et al. (2008) but for the stimuli used which were related to Luupite and Niffite. We collected response latencies for the two 72-trial critical blocks. In one of the critical blocks, participants categorized Luupite and self-related words with a single response key (e.g., the E key) and Niffite and others-related words with another single response key (e.g., the I key). In the other critical block, participants categorized Niffite and self-related words with a single response key and Luupite

and others-related words with another single response key. Order of these blocks was counterbalanced across participants.

After this task, participants were asked for their demographics. They were then thanked and debriefed by the experimenter.

Results and Discussion

We used the same IAT score calculation as Kawakami et al. (2008): We excluded incorrect trials (7.25 % of the trials) and we recorded response times (RT) under 300 ms and above 2,000 ms to respectively 300 ms and 2,000 ms (0.95 % of the remaining trials). Then, we log-transformed RT and computed the difference between critical blocks so that a positive score indicates that the participant was faster in the “self-Luupite” than in the “others-Niffite” block. We use this score as dependent variable in our analyses.

To test the hypothesis that an approach training toward a group of stimuli led to faster RT in the block where these stimuli shared the same response key with “self”, we conducted a between-participant ANOVA. Between-participant predictors were the approached target (Luupite approach vs. Niffite approach) and the task used during the training (Joystick task vs. VAAST). We excluded one statistical outlier of this analysis because of a gap on their Cook’s d (Judd et al., 2017).¹⁸

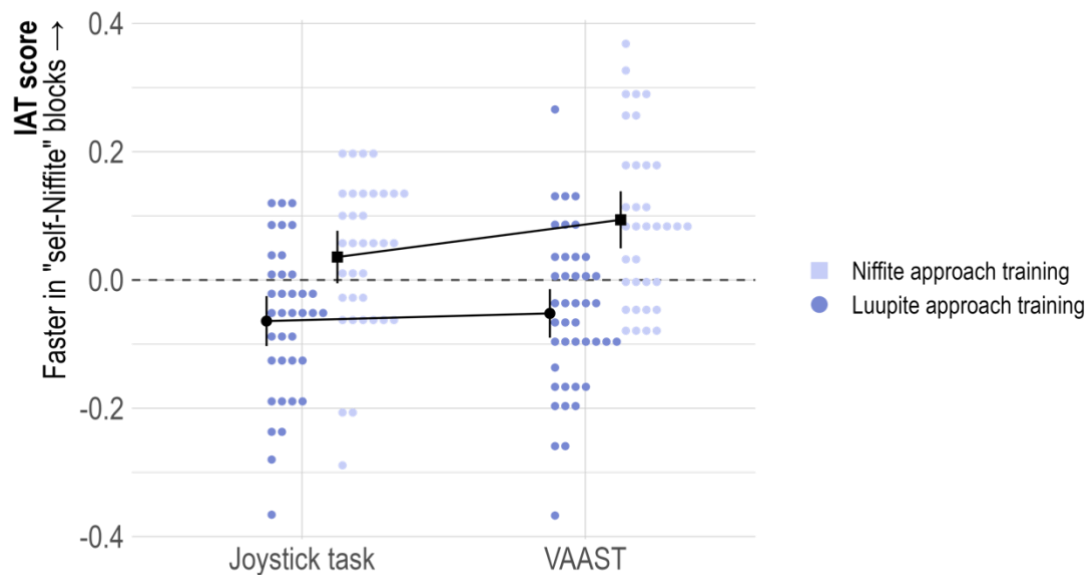
As predicted, this analysis revealed that the identification IAT difference scores were significantly lower when participants approached Luupite words than when they approached Niffite words, $t(145) = 6.14$, $p < .001$, 95% CI [0.083; 0.163], $\eta_p^2 = .201$ (see Figure 6.2). This effect indicates that participants were significantly faster in the block where the self and the

¹⁸ Not removing the statistical outlier neither change the direction nor the significance of the results.

list of words they approached were associated with the same response key. Additionally, the difference between the two tasks, descriptively in the direction of a superiority of the VAAST, was not significant, $t(145) = 1.15$, $p = .252$, 95% CI [-0.33; 0.125], $\eta_p^2 = .009$. Simple slope analysis revealed large movement effect for both the Joystick task and the VAAST, respectively, $t(145) = 3.33$, $p = .001$, 95% CI [0.043; 0.157], $d_s = 0.85$, and, $t(145) = 5.50$, $p < .001$, 95% CI [0.091; 0.201], $d_s = 1.18$. The main effect of the task used during the approach training was not significant, $t(145) = 1.75$, $p = .083$, 95% CI [-0.33; 0.125], $\eta_p^2 = .020$.

Figure 6.2

IAT Difference Score According to the Approach Training Condition and the Task Used for the Training



Note. IAT difference scores were computed by subtracting log-transformed RT for the block where “self” and “Luupite” were associated with the same response key to log-transformed RT for the block where “self” and “Niffite” were associated with the same response key. Each point represents the score for a given participant. Participants above the dotted horizontal line are faster in the “Niffite-self” block, participants below are faster in the “Luupite-self” block. Error bars correspond to 95% confidence interval. RT = response times.

To avoid relying on a null effect, we preregistered that the difference between the two tasks could be considered as negligible. Wellek (2010, pp. 278–284) provides a procedure to conduct negligibility tests of interactions. In this procedure, one needs to set an a priori limit the estimated effect size the interaction is not supposed to exceed for the effect to be considered negligible. The preregistered value for this limit was $\eta_p^2 = .013$. Because our

estimate effect size was $\eta_p^2 = .009$, the difference of effect between the tasks, in addition to not being significant, is negligible.

These results indicate that approaching and avoiding novel groups of stimuli influence their indirect evaluation: Participants are faster during the identification IAT when the self category and the approached list are associated with the same response key compared to when the self category and the avoided list are associated with the same response key. Given that the Joystick task effect was not significantly larger than the VAAST effect (and in fact descriptively smaller), because the VAAST comes with no ambiguity in response mapping, and because using a different task could increase generalizability, we used this task to conduct our Kawakami et al.'s (2008) replication.

Experiment 1b

As a second step in the validation of the VAAST approach and avoidance training, we decided to replicate findings of Experiment 1a but online. The recent credibility revolution brought a lot of changes aiming at increasing the robustness of psychological science (Vazire, 2018) and one of them was the use of online samples (rather than lab samples; Sassenberg & Ditrich, 2019). The use of online samples allows researchers to overcome some of the recently highlighted problems in psychology, namely, the representativeness of the participant samples and the size of the small samples recruited.

Concerns regarding the use of convenience sample such as college students have been expressed for a while now (e.g., Sears, 1986). Academic research often takes place in the universities where college student represents an accessible and convenient labor force, and, accordingly, these students are recruited as a non-negligible part of the participants in psychology experiments. Because students might be outliers in the general population (see also Henrich et al., 2010), recruiting participants outside of these convenience samples is a first step for increasing the robustness of findings in social psychology. Second, the use of

small sample sizes has been affecting the replicability of psychological science. When it comes to effect size estimation, small sample sizes come with more uncertainty (Schönbrodt & Perugini, 2013). This uncertainty, when combined with publication bias (C. J. Ferguson & Heene, 2012), results in the inflation of the number of false positive in the literature (Ioannidis, 2008). In the end, because online samples offer diverse samples and because running an experiment reduce the labor cost by making participants engage in the study on their own, developing online version of paradigms represents a necessary step to increase the robustness of the findings in psychology.

Contrary to the Joystick task, which would represent a logistical nightmare from the perspective of participants engaging with online experiments, the VAAST is a good candidate for an online paradigm. The VAAST does not require sophisticated apparatus other than a keyboard and a monitor. In fact, Aubé et al. (2019) successfully used it in four online experiments where they used the paradigm as a measure of approach-avoidance tendencies (see also Rougier et al., 2020).¹⁹ However, their implementation lacked some critical features regarding our use of the VAAST as an approach-avoidance training.²⁰ We decided to build

¹⁹ In their Experiment 3, Rougier et al. (2020) used an online version of the VAAST to see whether approaching/avoiding Black American names was easier than White American names. The version used was the same as the one that we developed for Experiment 1b.

²⁰ Aubé et al. (2019) used the psytoolkit framework to develop their online version of the VAAST (Stoet, 2010, 2017). While psytoolkit is an easy platform to work with, as it does not need advanced coding skills, it lacks flexibility we needed for our implementation. Some features of the VAAST are impossible to implement with the psytoolkit framework (e.g., several steps in the environment, error feedback).

our implementation of the VAAST from scratch using jsPsych (de Leeuw, 2015; de Leeuw & Motz, 2016), and to test it by replicating the VAAST condition of Experiment 1a.

In Experiment 1b, we replicated the Experiment 1a's VAAST condition online. Participants performed an approach and avoidance training with Luupite and Niffites stimuli (Gregg et al., 2006). Then, we assessed whether participants formed any evaluative responses towards these groups using an identification IAT (Nosek et al., 2002). We will highlight any difference with the Experiment 1a's VAAST condition in the following section. We will also highlight some specificities with online data collection. We decided to use the Prolific platform as it was designed to overcome some of the ethical limit of competing platforms such as MTurk (Fort et al., 2011; Palan & Schitter, 2018; Williamson, 2016). We report all measures, manipulations, and exclusions. Material, data, the analysis script, and the preregistration can be found online at: <https://osf.io/jv5ns/>.

Method

Participant, design, and sensitivity analysis

Eighty participants ($M_{\text{age}} = 35.40$, $SD_{\text{age}} = 11.10$; 36% female, 63% male, 1% other) took part in this study through the Prolific platform (Palan & Schitter, 2018). Participants took part in the study in exchange for £1.70. They were screened to be either British or American and to be native English speakers. We relied on a 2 (training: “Luupite approach” vs. “Niffite avoidance”) between-participant design.

Based on the results of Experiment 1a, we could expect the effect size of the approach and avoidance training on the identification IAT to be $d_s = 1.18$, 95% CI [0.69, 1.66].

Sensitivity analysis revealed we had 80% statistical power to detect an effect size of $d_s = 0.64$ with a sample size of 80 participants. With a conservative exclusion rate of 30%, we still had 80% statistical power to detect an effect size of $d_s = 0.77$. Accordingly, we consider that recruiting 80 participants made this study adequately powered.

Procedure and Material

Participants took part in the experiment within their web browser. They were recruited using the Prolific platform (Palan & Schitter, 2018). Upon accepting to take part in this study, they were redirected on the website where we hosted the experiment. Once the experiment started, it was automatically set on Fullscreen for participants to focus entirely on the study.

As in Experiment 1a, participants were randomly assigned to either the approach Luupite condition or the approach Niffite condition (see Gregg et al., 2006). Participants' task for the first part of the experiment was to approach the words of the Luupite list and to avoid the words of the Niffite list (or the opposite). As in Experiment 1a's VAAST condition, each trial started with a fixation cross for a random duration (between 650 ms and 1200 ms) followed by a target word at the center of the screen. When the word was presented, participants pressed the key corresponding to the movement they had to perform. Participants had to press either the up arrow or the down arrow key twice to complete a trial, each key press triggering an animation as if they were walking two steps forward/backward within the environment. If participants made an error, a red cross appeared on the screen and they had to start the trial over again. Participants first completed a 20-trial practice block and then moved to a 200-trial main block; one half of the trials was related to approach, and the other one was related to avoidance. The approach-avoidance training phase lasted approximately 10 minutes.

As in Experiment 1a, immediately after the VAAST approach-avoidance training phase, participants completed an identification IAT (see Nosek et al., 2002). In this task, participants categorized words from the Luupite and Niffite lists, and words related to the self (e.g., I, mine) and others (e.g., they, theirs; see Appendix 3). Unlike in Experiment 1a, participants used their keyboard rather than a Chronos box during the IAT. Unlike Experiment 1a, we collected response latencies for the two 72-trial critical blocks, but also

the 20-trial critical block practice. In one of the critical blocks, participants categorized Luupite and self-related words with the same response key (e.g., the E key) and Niffite and others-related words with another response key (e.g., the I key). In the other critical block, participants categorized Niffite and self-related words with the same response key and Luupite and others-related words with another response key. Order of these blocks was counterbalanced across participants.

Results and Discussion

For this experiment, we used a different (but pre-registered) calculation for the IAT score than the one used in Experiment 1a.²¹ In Experiment 1a, we used the same IAT computation as Kawakami et al. (2008) so that our upcoming replication would be as close as possible to the original study. The goal of Experiment 1b, however, was to assess whether we could use a VAAST based training online. Accordingly, we used a scoring algorithm which had been used in online studies before: the D2 algorithm (e.g., Van Dessel et al., 2017; Greenwald et al., 2003). The D2 scoring includes every trial where the participants categorize target words along the two types of categories (including practice trials). It uses a cutoff of 400 ms and 10 000 ms on reaction times and errors are included in the analysis. Then, the algorithm scales reaction time in the two types of critical blocks at a participant level (i.e., in “self-Luupite” blocks, reaction times for a given participant are divided by the participant’s standard deviation of reaction times for the “self-Luupite” blocks; Greenwald et al., 2003). Finally, the algorithm computes the difference between the two critical types of blocks so that a positive score indicate that participants were faster for the “self-Luupite” blocks compared to the “self-Niffite” blocks. This score served as dependent variable in our analyses.

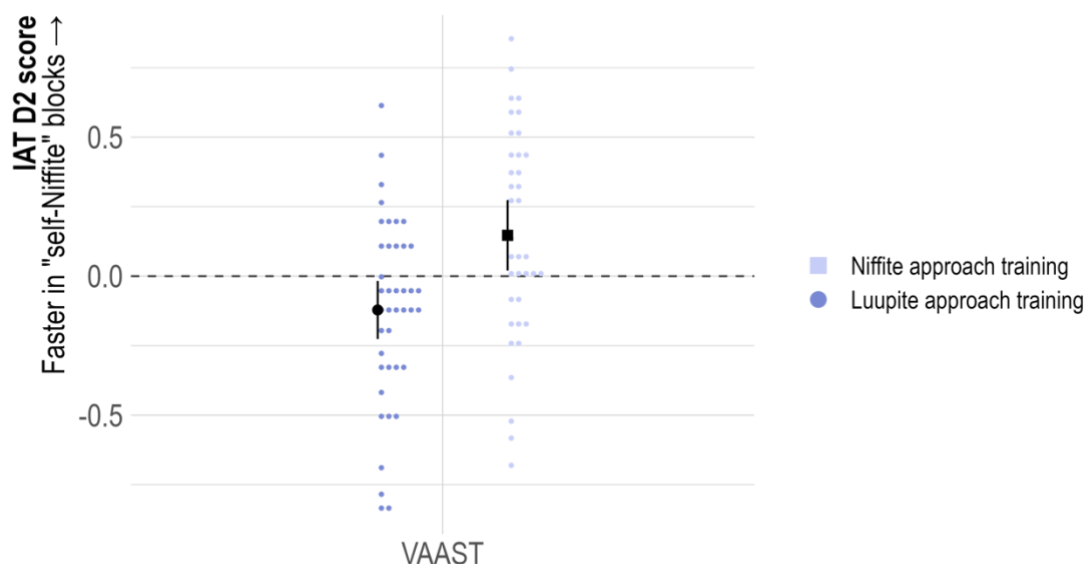
²¹ Adopting the same calculation as Experiment 1a does not change neither the direction nor the significance of our effects.

To test the hypothesis that an approach training toward a group of stimuli led to faster RT in the block where these stimuli shared a response key with “self”, we conducted a between-participant t -test with the approached group as a predictor. The recruitment of 80 participants on the Prolific gave us complete data for 78 participants. We preregistered to exclude outliers of this analysis following on Judd et al.'s (2017) recommendations. We did not exclude any participants based on these criteria.

Replicating results from the Experiment 1a, the analysis revealed a significant effect of the group to approach. Scores were significantly higher when participants had to approach the Luupite category rather than the Niffite category, $t(78) = 3.34$, $p = .001$, $d_s = 0.76$, 95% CI [0.29, 1.21] (Figure 6.3).

Figure 6.3

IAT Difference Score According to the Approach Training Condition



Note. Each dot represents a participant. Error bars represent confidence intervals. IAT scores were computed according to the D2 algorithm (Greenwald et al., 2003).

With Experiments 1a and 1b, we showed that the VAAST could produce learning effects competing with a procedure that is usually used in the literature (namely, the Joystick task; Experiment 1a) and that such a learning held with online sample (Experiment 1b). The acquisition of an indirect evaluative response appears to be reliable with a VAAST-based approach-avoidance training. With medium-to-large effect size obtained in the two experiments, we consider that the necessary conditions to be confident that the VAAST can produce evaluative response learning are satisfied and can try to replicate the work Kawakami et al. (2008) conducted.

Experiment 2

With Experiment 2, we aimed at replicating Kawakami et al.'s (2008) work on the effect of a math-approach training on an identification IAT among women with low identification to math. We expected participants' indirect evaluations of math to change after the training. Participants who approached math-related stimuli should be faster in the IAT block where math and self categories shared the same response key, compared to participants who avoided math-related stimuli. Note that, even if we were primarily interested in indirect evaluation change, we also included (at the end of the experiment) a direct evaluation of math identification for exploratory purposes. Material, data, the analysis script, and the preregistration can be found online at: <https://osf.io/amq99/>.

Method

Participant, design, and sensitivity analysis

Two hundred and three participants²² ($M_{\text{age}} = 20.91$, $SD_{\text{age}} = 3.37$) were recruited for this experiment. Participants took part in exchange for either class credit or 10€. We relied on a 2 (training: “math approach” vs. “math avoidance”) between-participant design.

Note that even if the original and replication context (respectively, Canada and France) are not identical, they are similar on several points. First, both Canada and France are WEIRD countries, that is Western, Educated, Industrialized, Rich, and Democratic (Henrich et al., 2010b). Moreover, in terms of gender gap in math achievement, OECD reports show quite similar numbers for Canada and France (OECD, 2015). These numbers indicate, for example, that boys outperform girls in problem solving in both countries, but also that the gender gap among top achievers in both Canada and in France is larger than the OECD average. We consider the context of replication as well as the sample used for the replication similar to the original Kawakami et al.’s (2008) paper.

Regarding statistical power, Funder et al. (2014) recommend adopting a more demanding threshold than the usual 80% recommended by Cohen (1988) when conducting a single study replication. Accordingly, we preregistered a 200 participants sample size for our replication as it allows us to reach a statistical power over 99% to detect an effect of Kawakami et al.’s (2008) meta-analytical effect size (i.e., $d_s = 0.64$). It is worth mentioning that one could argue that Kawakami et al.’s effect size might be inflated given the small sample sizes they adopted (Schönbrodt & Perugini, 2013). An alternative, safest, strategy could have been to ignore Kawakami et al.’s results and use the average effect size in social psychology instead ($d_s = 0.40$; Richard et al., 2003). A power analysis using this effect size

²² Data of two participants were lost due to a power outage during the procedure.

reveals 80% statistical power. Although this still complies with Cohen's (1988) recommendation, it can be considered suboptimal for a single study replication research (Funder et al., 2014).

Procedure and Material

We conducted the replication following Kawakami et al.'s (2008) Experiment 1 procedure. In what follows, we explicitly mention the difference between the original experiment and the replication. A summary of the difference between the Kawakami et al.'s original experiment and the replication can be found in the replication recipe in Appendix 2 (Brandt et al., 2014).

Participants were pre-screened by answering a questionnaire either online or in person. We recruited only female participants who reported not liking math (i.e., less than four on a "I like math" seven-point Likert item). Participants who fitted our inclusion criteria were invited to take part in the actual experiment. We informed participants that they would have to engage in a series of tasks aiming at testing material related to school affinity.

One out of four different experimenters welcomed the participants in the lab and led them in an individual cubicle where participants received instructions on how to complete the study. Participants were randomly assigned to one of the two experimental conditions—math approach condition or math avoidance condition—and the experimenter was blind to this condition.

As in Experiment 1a, participants were first told that they would have to approach and avoid several stimuli in the first part of the experiment. Participants were randomly assigned to either the math approach condition or the math avoidance condition. Participants in the math approach condition were instructed to approach math-related pictures ($N = 24$) and to avoid arts-related ones ($N = 24$); participants in the math avoidance condition had the opposite instructions. A subset of the stimuli used for the approach and avoidance training

can be found in Appendix 3. Note that the pictures we used differ from Kawakami et al.'s (2008) because we could not access original material. Therefore, we sampled 48 stimuli from Wikimedia. Participants were then informed on how to proceed.

Unlike Kawakami et al.'s (2008) experiment which adopted a Joystick task implementation of approach and avoidance, the training in our replication was similar to the VAAST condition of the Experiment 1a except for a few differences. Instead of words, participants approached or avoided pictures related to math or to arts. Instead of answering with keyboard directional buttons, participants had to press buttons on a *Chronos* response box. Approach and avoidance buttons were labeled with the "Approach" and "Avoid" labels. The total number of trials was 480 distributed in 10 blocks of 48 trials. One half of the trials was related to approach and the other one was related to avoidance. The training phase lasted approximately 25 minutes.

After the approach and avoidance training, participants completed an identification IAT with math and arts as target categories. The IAT procedure was the same as in the Experiment 1a. An exhaustive list of the stimuli used for the IAT can be found in Appendix 3.

A notable difference with Kawakami et al.'s (2008) original experiment is that we also decided to include a self-report measure assessing how much participants identify to math and arts. This measure is inspired by the Inclusion of Other in the Self scale (Aron et al., 1992), which was developed in the context of intergroup relationship and has already been used in the context of math identification (Necka et al., 2015). Participants were presented two circles and they could adjust how much the circles overlapped using the keyboard arrow keys. They were informed that one circle was representing themselves and the other was representing the concept of math (arts). Participants received the instruction to choose the overlap between the circles corresponding to the how much they thought their

sense of themselves was overlapping with the concept of math (arts). Participants answered for both math and arts concepts and the order was randomized across participants.

After this task, participants were asked for demographic information. They were then thanked and debriefed by the experimenter.

Analysis

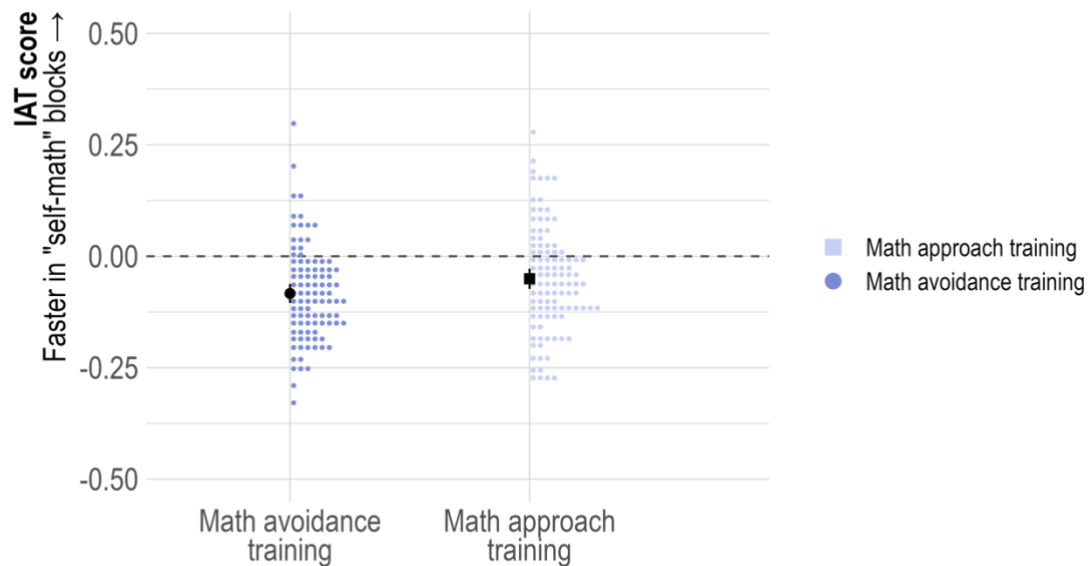
Identification IAT

As in Experiment 1a, we used the same IAT score calculation as Kawakami et al. (2008): We excluded incorrect trials (6.23% of the trials) and we recoded RT under 300 ms and above 2,000 ms (0.79% of the remaining trials) to respectively 300 ms and 2,000 ms. Then, we log-transformed RT and computed the difference between test trials on critical blocks so a positive score indicated that the participant was faster in the “self-math” than in the “self-arts” block. We used this score as a dependent variable.

We conducted a *t*-test comparing participants who approached math-related material and those who avoided math-related material. We found a significant difference between the two conditions indicating that participants in the approach math condition had higher scores ($M = -0.051$, $SD = 0.117$, $n = 104$) than participants in the avoid math condition ($M = -0.084$, $SD = 0.106$, $n = 97$), $t(199) = 2.10$, $p = .037$, 95% CI [0.002 ; 0.064], $d_s = 0.30$. This analysis revealed a small effect in the direction of our prediction (see Figure 6.4).

Figure 6.4

Identification IAT Difference Score for the Math approach Training and Math Avoidance Conditions



Note. IAT difference scores were computed by subtracting log-transformed RT for the block where “self” and “arts” were associated with the same response key to log-transformed RT for the block where “self” and “math” were associated with the same response key. Each dot represents a participant. Participants above the dotted horizontal line are faster in the “math-self” block, participants below are faster in the “arts-self” block. Error bars correspond to the 95% confidence intervals.

Self-overlap measure

We analyzed data from the self-overlap measure according to a mixed effect ANOVA with one between-participant predictor (i.e., training) and a within-participant predictor (i.e., target participants evaluated). We coded participants responses so that a score of 0 indicates

that the circles representing the self and the concept do not overlap and so a score of 100 indicates the two circles totally overlap.

This analysis revealed a main effect of the target, indicating that participants reported more self-concept overlap when the target was “arts” ($M = 62.63$, $SD = 25.46$, $n = 201$) than when the target was “math” ($M = 20.82$, $SD = 20.05$, $n = 201$), $t(199) = 16.26$, $p < .001$, 95% CI [36.83; 46.99], $\eta_p^2 = .569$. The main effect of the training was not significant, participants in the “math approach” condition ($M = 40.87$, $SD = 13.30$, $n = 104$) did not report higher overlap than participants in the “math avoidance” condition ($M = 42.64$, $SD = 14.38$, $n = 97$), $t(199) = 0.91$, $p = .364$, 95% CI [-5.63; 2.07], $\eta_p^2 < .001$. The interaction between the training condition and the target participants evaluated was not significant, $t(199) = 1.10$, $p = .271$, 95% CI [-15.86; 4.48], $\eta_p^2 = .002$: The difference between how participants overlapped with math and arts concept was not significantly reduced after a math approach training.

Complementary analyses

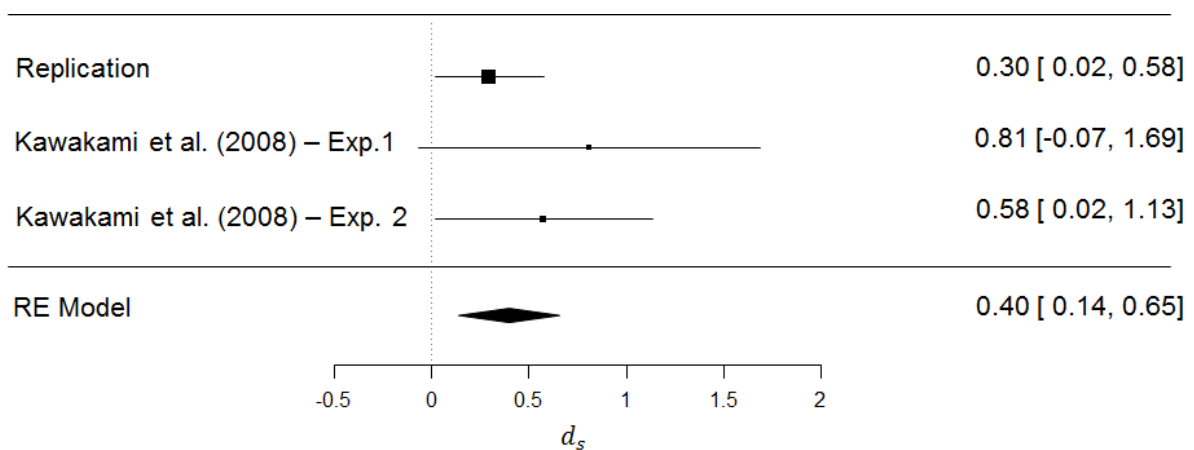
To better evaluate the contribution of the present experiment regarding Kawakami et al.’s (2008) original data on the effect of math approach training on math identification, we computed a Bayesian factor using the Savage-Dickey density-ratio method (Wagenmakers et al., 2010). A Savage-Dickey Bayes factor reveals how much our belief in a hypothesis changes after collecting data. This analysis revealed that the hypothesis of a null effect of the math approach and avoidance training is now 2.57 less. By computing a meta-analytic Bayes factor, we estimated that the current level of evidence for an effect of a math approach training compared to no effect is now “strong” ($BF_{10} = 18.77$; Jeffreys, 1961; Morey & Rouder, 2018).

Moreover, to summarize the current evidence for a math approach and avoidance training effect on math implicit identification among women with low identification to math,

we conducted a meta-analysis. We included Experiments 1 and 2 of Kawakami et al. (2008), as well as our replication for a total of 3 effects (i.e., $k = 3$) and 281 participants. We conducted a random-effects meta-analysis with the metafor package (REML estimation; Viechtbauer, 2010). The estimated effect size of the approach and avoidance training effect was $d_s = 0.40$, 95% CI [0.14; 0.65] and there was no evidence for variation in the effect sizes between studies, $Q(2) = 1.74$, $p = .419$. This meta-analysis suggests a small-to-medium effect size for the effect of approaching math-related material (instead of avoiding it) on an identification IAT (see Figure 6.5).

Figure 6.5

Meta-analysis of the Effect of Math Approach Training on an Identification IAT



Note. Random-effect meta-analysis on the math approach and avoidance training effect on a math identification IAT. Effect sizes have been coded so that positive d_s indicates an increase in math identification after a math approach training.

General Discussion

Reducing the gender gap related to math is still a major challenge in our society and implicit attitude change offers an interesting perspective on this issue (Nosek et al., 2009).

Kawakami et al. (2008) observed that women who repeatedly approached math-related material self-identified more to math (as measured with an identification IAT; Nosek et al., 2002). This original experiment, however, suffered from a quite modest evidential value. With the present replication, we aimed at providing stronger evidence in favor of the math approach training effect, as well as increasing the generalizability of Kawakami et al.'s (2008) finding. We conducted this replication with a sample size large enough to detect small-to-medium effect size and we preregistered our hypotheses and analytic strategy. Using a VAAST-based procedure (Rougier et al., 2018), a procedure that we wanted to validate, we replicated Kawakami et al.'s (2008) results, providing stronger evidence for an effect of math approach training (i.e., increasing the evidential value by a factor of 2.57). Building upon Kawakami et al. (2008) and endorsing today's research standards (see Vazire, 2018), this replication therefore contributes to providing experimental evidence that an approach and avoidance training can change evaluation.

Beside the results on the identification IAT, we decided to explore the impact of the approach and avoidance training on a more direct measure of identification (i.e., the self-overlap measure). Several reasons might explain the absence of significant difference between participants who approached math and those who avoided math. First, the statistical noise of a single-item measure makes the comparison between the conditions on the self-overlap measure inherently less powerful than the comparison on the IAT. However, recent results in the literature suggest that one can acquire a new indirect evaluative response, without having an impact on a more direct one. Indeed, Van Dessel et al. (2016) obtained such pattern of results in an experiment focusing on the effect of approach and avoidance instructions. Still, at this stage, one can probably infer that approach and avoidance training might not be the easiest way to change self-reported identification to math.

While Kawakami et al. (2008) relied on the idea that the mere association between a movement and a stimulus has an effect on implicit identification, as discussed in the introduction section of this dissertation, recent theoretical accounts have suggested propositional processes as a mechanism to explain the effect of approach and avoidance paradigm. According to Van Dessel et al. (2018a), performing an action influences evaluations because participants infer a relation between the stimulus and evaluation thanks to the available information during the approach-avoidance training. That is, when participants have to repeatedly perform an approach movement faced to a stimulus, the participants infer an association between the stimulus and the approach concept. Given that participants know that they usually approach things that are relevant to them, the approached stimulus would be tied to the self. Critically for the inferential account, when participants can infer the same information out of instructions they would have from the training, they do acquire the same evaluative response (Van Dessel, De Houwer, Gast, et al., 2016; Van Dessel et al., 2015).

One could thus argue that an actual training would not be the most efficient way to change implicit identification as instructions alone have an effect on indirect evaluations. The goal of the present dissertation is to provide empirical data to investigate this question. Nevertheless, Van Dessel et al. (2015) failed at showing an effect of the instructions for stimuli for which people already had strong attitudes (i.e., ethnic groups). Moreover, Van Dessel et al. (2020) recently showed that, when it comes to real social groups (i.e., Turkish and Flemish groups), actual training was an effective way to change indirect evaluative responses while instruction was not. Our results are in line with the literature suggesting that actual experience in interventions aiming at changing implicit identification possesses some critical features that are worth investigating when studying the change of strong evaluative

responses. The current replication is therefore an important contribution that confirms the robustness and validity of such interventions.

Regarding the effectiveness of the approach avoidance training on implicit identification to math, the overall effect size estimated from Kawakami et al.'s (2008) experiments and the current replication corresponds to a small-to-medium effect. Despite the differences between Kawakami et al.'s (2008) original work and our replication in terms of sample (Canadian vs. French), time (2008 vs. 2017), paradigm (Joystick task vs. the VAAST) and material (the pictures used in the training task), we retrieved a math approach training effect in our replication. This suggests generalizability of the approach training as a way to change math implicit identification. Moreover, this effect size is in line with the literature interested in interventions aiming at changing indirect evaluation. Lai et al. (2014) conducted a large-scaled study where they compared the effectiveness of different interventions aiming at reducing implicit racial preference. The most effective interventions could change implicit preference with effect sizes ranging up to $d_s = 0.49$. In comparison, the estimated math approach and avoidance training had an effect size of $d_s = 0.40$. The effectiveness of the approach and avoidance training intervention is, thus, close to the most effective interventions included in Lai et al. (2014) comparison. This highlights the importance to consider approach and avoidance training when it comes to changing implicit identification.

One limitation to mention, however, as to do with the sample size we used in this replication. As we discussed in our power analysis section, we performed our power analysis by relying on the effect sizes observed in Kawakami et al. (2008)'s work, but (their) small sample sizes also come with imprecise effect sizes. Hence, our study can be seen as underpowered (i.e., 80% power instead of 90%, although this could be somewhat compensated by using a conservative two-tailed test) if one instead uses the average effect size in social psychology (Richard et al., 2003).

Conclusion

In this chapter, with a highly powered close replication, we showed that approach and avoidance training can change indirect evaluation, successfully replicating Kawakami et al.'s (2008) results. While endorsing the latest standards in terms of research practices (i.e., preregistration, open material, open data), we showed that the VAAST (Rougier et al., 2018) can be a viable implementation of approach and avoidance movement in approach-avoidance trainings. By using it rather than the traditional Joystick task, we increased generalizability of Kawakami et al.'s (2008) findings and updated the level of evidence for a math approach training effect (compared to no effect) by a factor of 2.57. With the experimental evidence gathered in this chapter validating the VAAST as an effective approach-avoidance training procedure, we will be able to move on to this dissertation research question, that is estimating the contribution of repeated experience in the effects of approach-avoidance training.

Chapter 7

Comparing Approach-Avoidance Procedures

When new tools emerged in the implicit social cognition literature (see Payne & Gawronski, 2010), some psychologists realized that these tools could overcome some of the limitations of self-report measures. In a nutshell, at first, these tools were thought to capture something so deeply rooted in the human mind that they could overcome the aforementioned problems of self-report measures. Because of this initial formulation, modern theories often assume that the acquisition of a specific response on those new tools—an indirect evaluative response—was the result of something repeatedly experienced. Recent research, however, shows that verbal statements (e.g., mere approach-avoidance instructions; Van Dessel et al., 2015) suffice for new indirect evaluative responses to emerge. Now, one question remains, is it the same to learn with repeated experience than it is to learn with mere instructions? We decided to address this question in the present chapter.

Repeated Experience and the Acquisition of Indirect Evaluative Response

As discussed in the introduction section of this manuscript, several theoretical models accounting for evaluation coexist. These models notably differ in how much they assume associative processes to be responsible for indirect evaluative response acquisition (see the Associative Processes section). The SEM (McConnell & Rydell, 2014), for example, make the assumption that associative processes are the only processes responsible for the acquisition of an indirect evaluative response. Because the core idea behind these associative processes is the idea that repeated co-occurrences tie concepts together, this model assumes a superiority of repetition-based paradigms over instruction-based paradigms when it comes to the acquisition of indirect evaluative response. In fact, the SEM assumes that merely

receiving instructions could result in the acquisition of direct evaluative responses, but not in the acquisition of indirect ones. The APE model and single-process propositional models do not make this assumption (De Houwer, 2018; Gawronski & Bodenhausen, 2007; Van Dessel et al., 2018a).

Both the APE model and single-process propositional models of evaluation suggest that indirect evaluative response can emerge through propositional processes. Rather than describing mere association emergence as a result of repeated co-occurrence, propositional processes describe how conditioned associations between concept—such as “I like math” or “Léa is Tom’s niece” (see Propositional Processes section)—emerge as a result of a process often seen as attention-consuming. In a theoretical model, Van Dessel et al. (2018a) explain how approach-avoidance training can result in the acquisition of an indirect evaluative response. In approach-avoidance training, Van Dessel et al. (2018a) assume that participants make several inferences. First, they are approaching (avoiding) a stimulus. Second, what they approach (avoid) is usually positive (negative) to them. Since they are approaching (avoiding) the stimulus, it must be positive (negative). While single-process propositional models assume that these propositions directly influence indirect evaluative response (e.g., De Houwer, 2018), the APE model suggests that the acquisition of the propositions can cause the emergence of a mere association between approach (avoidance) concept and the stimuli, and this mere association will then be captured by indirect evaluative response (Gawronski & Bodenhausen, 2006, 2007, 2014b, 2018). As long as one can make the relevant inference, propositional processes predict that both approach-avoidance training and approach-avoidance instructions can result in the acquisition of an indirect evaluative response. Note that both the APE model and single-process propositional models make predictions regarding the situations in which we should observe a larger difference between repeated-experience- and instructions-based procedures.

Indeed, the predictions regarding the difference between the two procedures are more nuanced in these models than it is within the SEM (McConnell & Rydell, 2014). First, the APE model assumes the dominance of a direct route to indirect evaluation passing by associative processes over an indirect route where propositional processes impact associative ones which, in turn, impact indirect evaluative responses. This superiority should be especially observable in situations where propositional and associative processes should result in opposite evaluation. As a result, in procedures where participants first acquire an evaluative response, and then acquired a second one in an opposite direct (sometimes referred to as counter-conditioning procedures), actual repeated experience outperforms mere instructions in changing the evaluation acquired through experience (Hu et al., 2017). Second, in their model describing the processes underlying the acquisition of evaluation resulting from an approach-avoidance procedure, Van Dessel et al. (2018a) assume that the repetition of an experience can increase the likelihood to infer the proposition associating a stimulus to a movement. As a consequence, assuming a situation without a ceiling effect where the relevant inference is easy to make, the number of trials in an approach-avoidance training should moderate evaluation acquisition (e.g., Woud et al., 2011). In the end, every model reviewed consider the possibility that approach-avoidance training, compared to approach-avoidance instructions, could result in the acquisition of a more pronounced evaluative response.

Approach-Avoidance Procedures

In the present Chapter, we decided to use approach-avoidance procedures to investigate the differences between repetition- and instructions-based procedures when it comes to indirect evaluative response acquisition. One of the benefits in adopting these paradigms is that we already know that both approach-avoidance actual training and approach-avoidance instructions result in the acquisition of indirect evaluative response

(Batailler et al., 2021; Kawakami et al., 2008; Van Dessel et al., 2015). Moreover, our lab developed the VAAST, a paradigm implementing approach and avoidance movement in a computerized paradigm (Rougier et al., 2018).

Approach-avoidance training implement the principle of repeated experience by making participants approach and avoid stimuli over several trials. In the VAAST, the procedure that we used as an approach-avoidance training in the **Chapter 6** (Experiments 1a–2), participants interact with stimuli in a computerized task. In the VAAST-based training, participants repeatedly approach one class of stimuli and avoid another one by moving in an environment. Usually, approach-avoidance training begins with an instruction phase where participants are familiarized with the stimuli and are told which stimuli they would have to approach and which stimuli they would have to avoid. Approach-avoidance instructions-based paradigms differ from training-based one in the moment the intervention stops. In approach-avoidance instructions procedures, participants receive the approach-avoidance instructions, and then, researchers usually investigate whether the evaluations of the stimuli changed (e.g., Van Dessel et al., 2015). By implementing the two procedures, we could evaluate the contribution of repeated experience regarding indirect evaluative response acquisition.

In Experiments 1a–2, we focused on the effect of approach-avoidance training on the acquisition of an evaluative response on an IAT. In Experiments 3 and 4, comparing the effects of approach-avoidance training and approach-avoidance instructions, we decided to move away from the IAT to focus on another indirect evaluative response: the Affect Misattribution Procedure (AMP; Payne et al., 2005). In the next section, we will detail some of the advantages of the AMP.

Capturing Evaluation with the Affect Misattribution Procedure

One of the main challenges when starting a line of research comparing repetition-based and instruction-based procedures was to be sure that we would be able to reliably observe the differences between the procedure. In other words, we had to be sure that the indirect evaluative response we would use had a good reliability. Many of these procedures exist in the literature: the AMP, the affective priming task, the Go-/No-Go Association Task, the IAT, to name a few (Fazio et al., 1995; Greenwald et al., 1998; Nosek & Banaji, 2001; Payne et al., 2005). With this chapter, we decided to use the AMP which adopts the same design as the evaluative priming task but which does not suffer from the low reliability that affects many indirect evaluative response measures (Gawronski & Hahn, 2019). Compared to the IAT, procedures like the evaluative priming task or the AMP adopts a design which makes it easier to disentangle the evaluation of the different stimuli.

The AMP (Payne et al., 2005) is one of the several paradigms whose design is similar to priming procedures such as the evaluative priming task (Fazio et al., 1986). This task is composed of several trials in which both a prime and a target are presented, and participants must report the valence which they associate with the target, ignoring the prime. Contrary to the evaluative priming task (Fazio et al., 1986, 1995), however, the AMP requires participants to report their subjective impression of the target. Rather than using positive and negative words as targets and determine whether the presentation of a prime interfere with valence recognition, AMP uses targets where, for most participants, there is no clear valence associated with the stimulus.²³ Since it was developed, this procedure was successfully used

²³ Target stimuli used in the AMP are Chinese ideograms, stimuli which would not be neutral for the 1.31 billion people speaking Chinese and its varieties. Still, as the majority of

in prejudice research (Imhoff & Banse, 2009), to investigate political behavior (Carraro et al., 2010), or even to predict substance use (Payne et al., 2008). For a review of the AMP, refer to Payne and Lundberg (2014).

In this chapter, in Experiments 3 and 4, we will compare the evaluative responses formed on an AMP between participants repeatedly experiencing the approach and avoidance movements toward a target and participants merely receiving approach-avoidance instructions. Experiment 3 will serve as a pilot experiment.

Experiment 3

In Experiment 3, we wanted to test a paradigm to see whether participants could acquire an evaluative response on an AMP (Payne et al., 2005) as a result of approach-avoidance procedures. The procedures that we included were an approach-avoidance training and approach-avoidance instructions. The goal of this experiment was primarily to test whether we could produce results consistent with the literature. We recruited participants enrolled in another experiment. In this other experiment, participants had to evaluate faces on several dimensions (see Courset et al., 2018). Experiment 3 always occurred after the other experiment. As a pilot study, Experiment 3 was not preregistered. We report all measures, manipulations, and exclusions. Material, data, and the analysis script can be found online at: <https://osf.io/nu9fg/>.²⁴

participants used in social psychology literature are coming from WEIRD countries (Henrich et al., 2010a, 2010b), this does not appear to be a problem regarding the validity of the task.

²⁴ This repository might be subject to an embargo. If it is unavailable, please e-mail cedric.batailler@gmail.com.

Method

Participant, Design, and Sensitivity Analysis

Sixty-three ($M_{age} = 22.39$, $SD_{age} = 3.58$, 57 women) psychology students from Université Grenoble Alpes (France) took part in this experiment in exchange for course credit. We did not screen our participant for this study.

In our experiment, participants could either be instructed to approach a logo or another (approach logo A vs. approach logo B), and they could be either assigned to the approach-avoidance training or the approach-avoidance instructions condition. The primes in the AMP were the (to be) approached logo and the (to be) avoided logo for every participant. We were interested in how the logos influenced the evaluation of a target.

Regarding the power of this analysis, with fewer than 50 participants by training condition in a between-participant design, this experiment can be considered as severely underpowered (Simmons et al., 2013). With a sample size of 63 and experimental design, we have .80 power to observe a difference of evaluative learning between the approach-avoidance training and approach-avoidance instructions if this effect is large ($d_s = 0.71$). The sensitivity analysis reveals a .12 power for a small effect size ($d_s = 0.20$), .50 power for a medium effect size ($d_s = 0.50$), and .88 power for a large effect size ($d_s = 0.80$; Cohen, 1988). Regarding the overall difference of evaluation between the (to be) approached and the (to be) avoided logos, sensitivity analysis revealed a .35 power to detect a small effect size ($d_z = 0.20$), .97 power to detect a medium effect size ($d_z = 0.50$), and over .99 power to detect a large effect size ($d_z = 0.80$). We had to adopt this sample size for reasons associated with the other study our participants were recruited for.

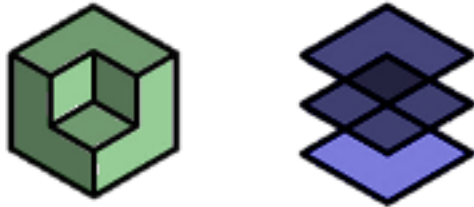
Procedure and Material

Participants took part in the experiment in the lab, in a room where up to eight participants at the time could complete the study. They were initially recruited to evaluate faces on several dimensions (see Courset et al., 2018); We used this as an opportunity to test whether approach-avoidance procedures were effective at forming evaluative responses on an AMP (Payne et al., 2005). Once participants completed the experiment for which they had been recruited, one out of two experimenters—blind to the condition of the participant—invited them to complete an unrelated experiment on attentional processes and information processing (Experiment 3). This experiment was coded using E-Prime.

For this experiment, we designed two logos. Participants were first informed that they would be exposed to these two logos during the experiment (see Figure 7.1). The procedure then differs according to the condition participants had randomly been assigned to (approach-avoidance training vs. approach-avoidance instructions). Conceptually speaking, the experiment could be seen as segmented in several blocks. These blocks were the approach and avoidance instructions, the approach and avoidance training, and the AMP. While participants in the approach-avoidance training condition received approach and avoidance instructions, completed an approach and avoidance training, and completed an AMP, participants in the approach-avoidance instructions condition simply received approach and avoidance instructions and completed the AMP (see Figure 7.2).

Figure 7.1

Logo Used as Stimuli in Experiment 3



Note. Left = Logo A, Right = Logo B

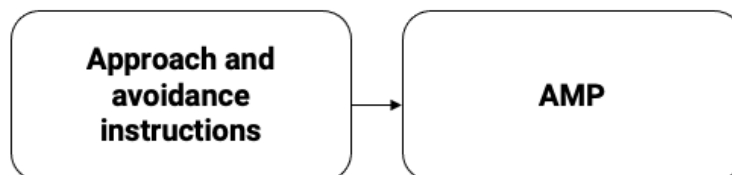
Figure 7.2

Time Course of Experiment 3 According to the Experimental Condition

Approach and avoidance training condition



Approach and avoidance instructions condition

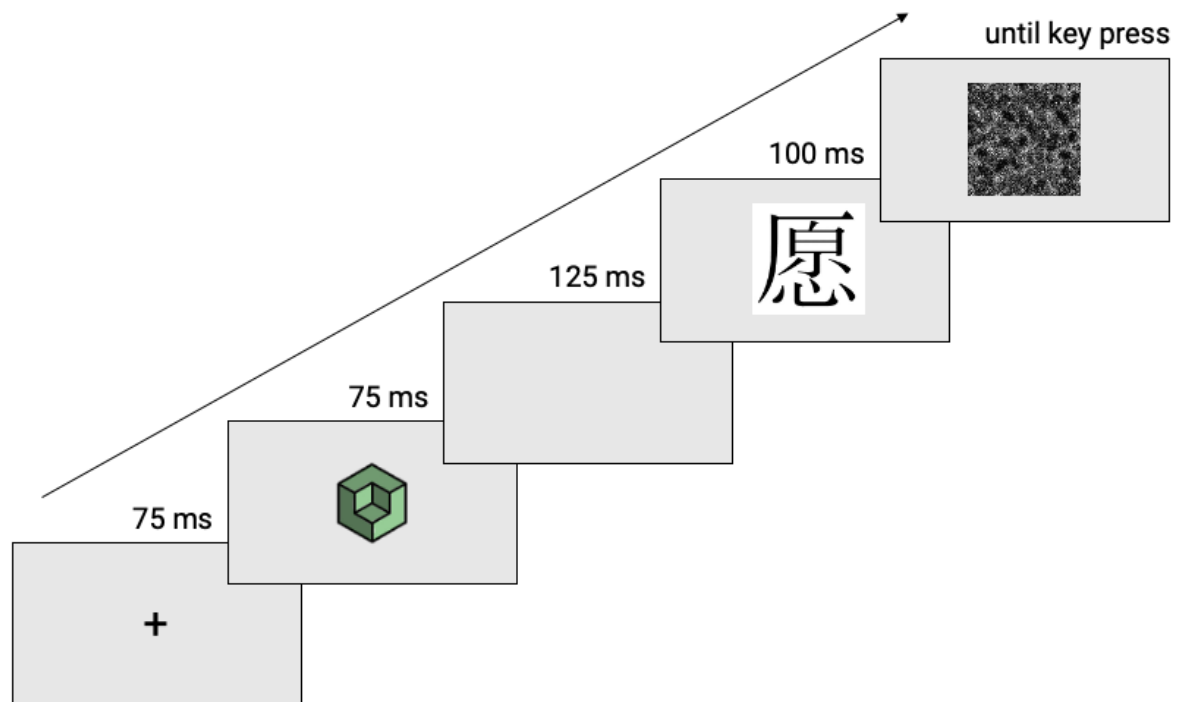


Approach-Avoidance Instructions. We adapted the instructions from an existing approach-avoidance instructions procedure (Van Dessel, Gawronski, et al., 2017). After being introduced to the logos (see Figure 7.1), participants were informed that the experiment was segmented in several tasks. One of them involved approaching and avoiding the logos. They were specifically instructed that they would have to approach the logo A and avoid the logo B (or the opposite, according to the condition they had been randomly assigned to). Because instructions would not be presented again during the experiment, the fact that participants had to remember them was emphasized (the verbatim we used was translated from Van Dessel, Gawronski, et al., 2017). Participants were told that we would explain to them later in the experiment how they would have to perform the approach and avoidance movement. Depending on the condition of the participant (i.e., approach-avoidance training vs. approach-avoidance instructions), participants would then move to the training or to the AMP.

Approach-Avoidance Training. The participants in the approach-avoidance training condition performed the training after receiving the instructions. This procedure was very similar to the one we developed in Experiment 1a. Participants performed the training in a virtual environment where they used the arrow keys of the directional keypad to move. Each trial started with a fixation cross for a random duration (between 650 ms and 1200 ms) followed by one of the two logos. Participants then had to press the correct key twice to complete the trial, each key press triggering an animation. Errors resulted in feedback asking participants to complete the trial again. The VAAST-based approach-avoidance training contained a practice block of 10 trials followed by five test blocks of 20 trials (for a total of 110 trials).

AMP. The evaluative response measure we used was an AMP (Payne et al., 2005). This task was introduced to participants as a way to understand how individuals do simple

but quick judgments. The task of the participants was to evaluate whether the Chinese ideograms were pleasant or unpleasant, compared to an average Chinese ideogram. Two stimuli were sequentially presented; Participants had to ignore the first one and evaluate the second one. For each trial, participants therefore saw a fixation cross for 75 ms, followed by the stimulus they had to ignore for 75 ms (i.e., the prime), a blank screen for 125 ms, the stimulus they had to evaluate for 100 ms (i.e., the target), and a mask until they answered (see Figure 7.3). Participants could either press the “I” keyboard key to state that they considered the stimulus pleasant or the “E” keyboard key to state that they considered it unpleasant. Critically, for the present experiment, the prime could be either a (to be) approached or a (to be) avoided logo. The AMP contained 4 practice trials, followed by 32 test trials. Reliability of the AMP was computed using the same calculation as in Payne et al. (2005) and reached $\alpha = .65$, 95% CI [.53; .78]. The reliability of the AMP we used was in the range of what is usually observed, [.47; .95] (Payne & Lundberg, 2014).

Figure 7.3*Time Course of an AMP Trial in Experiment 3*

At the end of the experiment, participants were debriefed, and thanked for their participation to this study. Participants in the approach-avoidance instructions conditions were told that they would not have to approach and avoid the logos. They committed themselves not to disclose the objectives of the experiment.

Results and Discussion

We wanted to test whether targets preceded by the (to be) approached primes with approach were evaluated more positively than the (to be) avoided primes. We also wanted to test whether this effect was moderated by the condition: Is the difference of effect stronger for participants in the approach-avoidance training than it is for participants in the approach-avoidance instruction condition? Because features of the logo could influence responses in

the AMP (participants could prefer the blue logo compared to the green one), we also included the logo in our analysis (approach logo A or approach logo B).²⁵

Analytically speaking, we used *lme4* to conduct a binomial mixed model (Bates, Mächler, et al., 2015). In this model, we were interested in the factors making participants more likely to answer “pleasant” than “unpleasant” during the AMP. Fixed effects predictors were the movement associated the prime was associated with (approach vs. avoidance), the training condition (approach-avoidance training vs. approach-avoidance instructions), and the logo used as a prime (logo A vs. logo B). Participants were used as a random factor in our model (for a discussion regarding the use of the terms random and fixed, see Nalborczyk et al., 2019). As recommended by Barr et al. (2013), we used the maximal random-effect structure for this model (see Appendix 4). Data from the 63 participants performing several test trials of the AMP made a total of 2,016 observations included in the analysis. We did not pre-register this analysis. For readability reason, we only interpret the effects that are relevant to the theory in the main text, but a summary of every coefficient can be found in Table 7.1. Appendix 5 contains a figure of the results per logo.

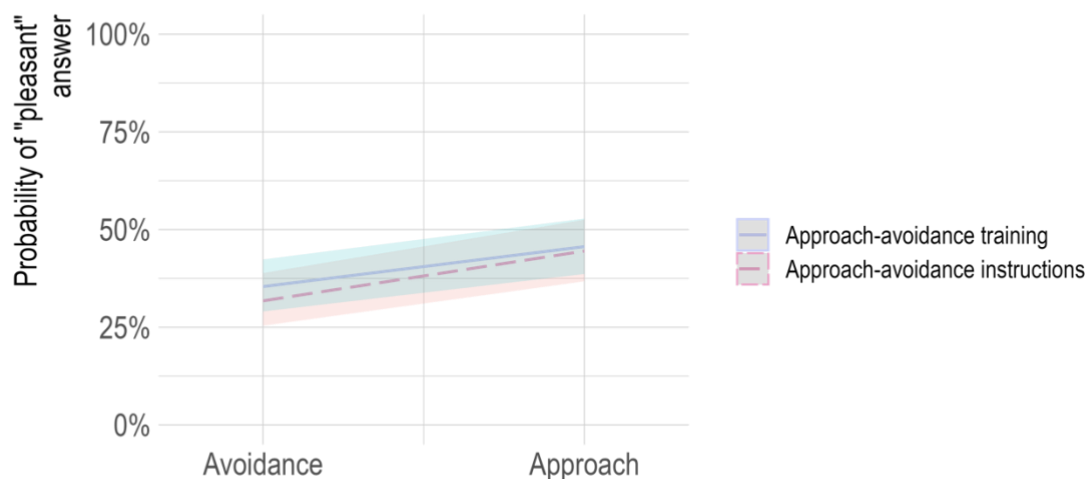
First, the mixed model analysis revealed that participants were overall less likely to find an ideogram pleasant than unpleasant, $z = 6.64$, $p < .001$, $OR = 0.64$, 95% CI [0.57; 0.73]. The main effect of the movement associated with the prime was significant, participants were 1.62 more likely to find an ideogram pleasant when it was associated with approach compared to avoidance, $z = 2.77$, $p = .006$, 95% CI [1.15; 2.29]. This analysis did not reveal a significant main effect of the condition of participants, $z = 0.79$, $p = .428$ $OR = 1.11$, 95% CI [0.86; 1.44]. Finally, the interaction between the movement

²⁵ Not including this logo in our analysis does not change the direction nor the significance of our effects.

associated with the prime and the conditions: The probability difference to evaluate the target as pleasant between primes associated with approach and primes associated with avoidance was not significantly stronger for participants in the approach-avoidance training condition compared to participants in the approach-avoidance instruction condition, $z = 0.33$, $p = .735$, $OR = 0.89$, 95% CI [0.45; 1.76] (see Figure 7.4).

Figure 7.4

Probability of a “Pleasant” Answer During the AMP According to the Features of the Prime and Participants’ Training Condition in Experiment 3



Note. Confidence interval represents the 95%CI of prediction at the population level based on the statistical model. The predictions are for a mean effect of the logo (logo A vs. logo B).

Table 7.1*Summary Statistics of Experiment 3's Mixed Model*

	log(OR)	95% CI	z	p-value
(Intercept)	-0.44	-0.57, -0.31	-6.64	< .001
Movement	0.49	0.14, 0.83	2.77	.006
Training	0.10	-0.15, 0.36	0.79	.428
Stimulus	-0.07	-0.41, 0.28	- 0.38	.707
Movement × Training	-0.12	-0.80, 0.57	- 0.34	.735
Movement × Stimulus	-0.17	-0.69, 0.35	- 0.63	.526
Training × Stimulus	-0.03	-0.71, 0.65	- 0.09	.931
Movement × Stimulus × Training	-1.3	-2.4, -0.30	-2.53	.011

Note. The Movement × Stimulus × Training effect can be visually examined in Appendix 5.

CI = Confidence Interval. OR = Odds Ratio.

With the results of Experiment 3, we provide initial evidence that logos associated with approach are evaluated more positively than logos associated with avoidance on an AMP. Given our sample size in this experiment ($n = 63$), the non-significance of the difference between the two approach-avoidance procedures (i.e., training and instructions) should not be seen as surprising. As indicated by the sensitivity analysis, with such a sample size, only large effect size could reliably be observed. By adapting the design and sample size, our next experiment will make it easier to observe any difference between approach-avoidance training and approach-avoidance instructions.

Experiment 4

As discussed in the Participant, Design, and Sensitivity Analysis section of Experiment 3, the test of difference of effect between approach-avoidance training and approach-avoidance instructions was severely underpowered (Simmons et al., 2013). To overcome this limitation, in Experiment 4 we decided to increase the sample size and to switch to a within-participant design for the approach-avoidance procedure condition (Perugini et al., 2018). Rather than having participants in the approach-avoidance training and participants in approach-avoidance instruction condition, every participant will be primed with both stimuli that they actually approached and avoided during a training and stimuli for which they merely received approach-avoidance instructions in the AMP. This means that, contrary to Experiment 3, participants will have to interact with four different logos. This experiment was preregistered. We report all measures, manipulations, and exclusions.

Material, data, the analysis script, and the preregistration can be found at

<https://osf.io/qty6w/>.²⁶

Method

Participant, Design, and Sensitivity Analysis

One hundred and three participants ($M_{age} = 20.60$, $SD_{age} = 3.04$, 81 women, 23 men) psychology students from Université Grenoble Alpes (France) took part in this experiment in exchange for course credit. Participants were screened out of the experiment if they participated in Experiment 3.

As in Experiment 3, our variable of interest was the evaluation of the target during the AMP. Our participants engaged in a 2-by-2 within-participant design. Primes presented

²⁶ This project might be subject to an embargo. If it is unavailable, please e-mail cedric.batailler@gmail.com to gain access to this project.

during the AMP could be associated with either approach or avoidance and they could have been actually approached and avoided during the training or not.

Our power analysis revealed that we had .51 statistical power to detect a small size effect of approach-avoidance training ($d_z = 0.20$), over .99 statistical power to detect a medium size effect of approach-avoidance training ($d_z = 0.50$), and over .99 statistical power to detect a large size effect of approach-avoidance training ($d_z = 0.80$; Cohen, 1988).

Compared to Experiment 3, as desired, the statistical power to detect the difference of effect between training compared to mere instructions was substantially increased. The decision to adopt this sample size was based on resource constraints and feasibility.

Procedure and Material

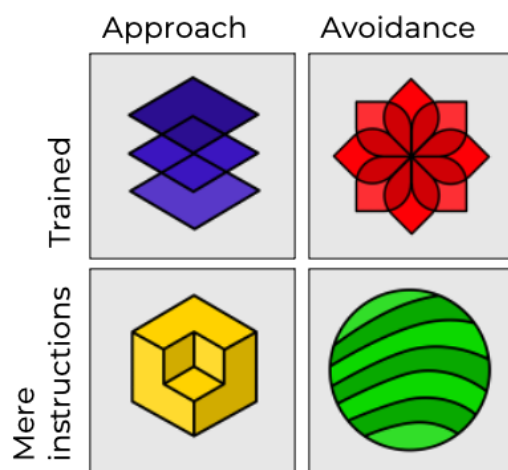
Participants took part in the experiment in the lab, in individual cubicles. This experiment was introduced as an experiment investigating attentional processes and information processing. Experiment 4 was coded using E-Prime. Before anything else, participants were first informed that they would be exposed to four different logos during the experiment. These logos were randomly generated so they all had a unique shape and a unique color (see Figure 7.5). They were generated based on a preset of four shapes and could be either blue, red, yellow, or green.

After being exposed to the logos, participants were informed that they would approach two of them and avoid the others. As in Experiment 3, we adapted the verbatim from Van Dessel et al. (2017) to tell participants that it was especially important that they remembered the movements associated with the logos. Following these instructions, every participant engaged with a VAAST-based approach-avoidance training similar to the one used in Experiment 3. We manipulated the experience with the stimuli (training vs. no training) by presenting only two logos in the training: one approached and one avoided. By the end of the training, participants had interacted with four stimuli: two associated with

approach and two associated with avoidance; two for which they actually had been trained and two for which they merely received instructions (see Figure 7.5). As in Experiment 3, the training contained a practice block of 10 trials followed by five test blocks of 20 trials (for a total of 110 trials). Participants were reminded of the movements associated with the logos right after the training and before moving to the AMP.

Figure 7.5

Example of Stimuli Used for Experiment 4



Note. For each type of the stimuli, a logo was randomly generated to have a unique shape and a unique color.

As in Experiment 3, we used an AMP to capture the evaluation of stimuli (Payne et al., 2005). Each trial of the AMP contained a prime, which was a stimulus presented in the experiment, and a target. Critically, the prime could be either associated with approach or with avoidance and participants could have either been trained to make a movement toward the prime or not. Because participants evaluated both stimuli for which they had been trained and stimuli for which they had not, the total number of trials increased compared to Experiment 3. The AMP contained 4 practice trials, followed by 68 test trials. Reliability of

the measure was computed using the same calculation as in Payne et al. (2005) and reached $\alpha = .87$, 95% CI [.78; .95].

After the AMP and before the end of the experiment, we asked participants to recall the movement associated with each of the logo. Indeed, awareness between a stimulus and its associated movement moderates the acquisition of evaluative responses following approach-avoidance procedures (Van Dessel, De Houwer, & Gast, 2016). Because we wanted to test the direct effect of repeated experience on evaluation acquisition, we pre-registered that we would exclude participants who could not recall the movement associated with a stimulus. To test whether participants could recall the movement associated with a stimulus, participants were shown the stimuli in a random order and asked to indicate whether they were associated with approach or avoidance.

At the end of the experiment, participants were debriefed, and thanked for their participation. They committed themselves not to disclose the objectives of the experiment.

Results

In order to be consistent regarding data analysis of the AMP, we deviated from the pre-registered analysis. Here, we report a logistic mixed model analysis instead of a by-participant ANOVA. This analysis is preferred compared to the pre-registered one because mixed modeling is considered more informative and less susceptible to type I errors (Judd et al., 2012). The reported analysis does not differ in terms of direction of the effects nor significance compared to the pre-registered one.

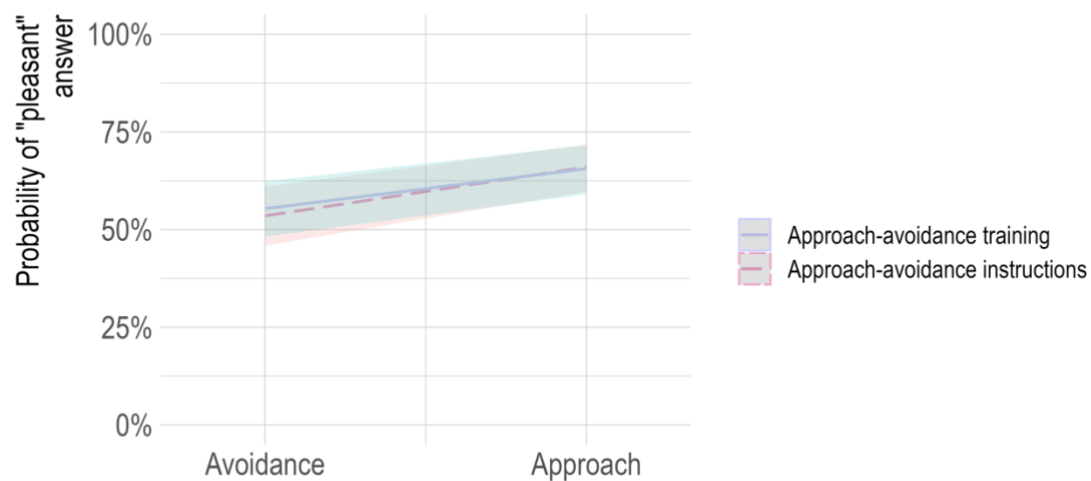
As in Experiment 3, we used *lme4* to conduct a binomial mixed model (Bates, Mächler, et al., 2015). In this model, we were interested in the factors making participants more likely to answer “pleasant” than “unpleasant” during the AMP. Fixed effect predictors were the movement was associated with the prime (approach vs. avoidance) and the experience of the participant with the prime (approach-avoidance training vs. mere approach-

avoidance instructions). Participants and stimuli were used as a random factor in our model. Because of convergence issues, we did not use the maximal random-effect structure (Barr et al., 2013), but followed the recommendations by Bates et al. (2015) to simplify the model (see Appendix 4). Data from the 103 participants responding (several times) to the 16 stimuli made a total of 7,004 observations included to fit the model. As pre-registered, we excluded data from one participant who could not remember the movements associated with the stimuli. We did not include practice trial.

Mixed model analysis revealed that participants were overall 1.52 more likely to find an ideogram pleasant than unpleasant, $z = 4.17$, $p < .001$, 95% CI [1.25; 1.84]. The main effect of the movement associated with the prime was significant, participants were 1.61 more likely to find an ideogram more pleasant when it was associated with approach compared to avoidance, $z = 2.95$, $p = .003$, 95% CI [1.17; 2.21]. This analysis did not reveal a significant effect of whether the prime was seen during the training or not, $z = 0.23$, $p = .821$, $OR = 1.03$, 95% CI [0.80; 1.32]. Finally, contrary to our hypothesis, the interaction between the movement associated with the prime and whether the prime was seen during the training was not significant. The difference of chance to evaluate an ideogram as pleasant when it is preceded with a prime associated with approach compared to a prime associated with avoidance movement is not significantly higher when the prime was seen during the training, $z = -0.79$, $p = .427$, $OR = 0.91$, 95% CI [0.73; 1.15]. Summary statistics can be found in Table 7.2.

Figure 7.6

Probability of a “Pleasant” Answer During the AMP According to the Features of the Prime in Experiment 4



Note. Confidence interval represents the 95%CI of prediction at the population level based on the statistical model.

Table 7.2

Summary Statistics of Experiment 4’s Mixed Model

Characteristic	log(OR)	95% CI	z	p-value
(Intercept)	0.42	0.22, 0.61	4.17	< .001
Movement	0.48	0.16, 0.79	2.95	.003
Training	0.03	-0.22, 0.27	0.23	.821
Movement × Training	-0.09	-0.32, 0.14	-0.79	.427

Note. OR = Odd Ratios, CI = Confidence Interval.

General Discussion

With Experiment 3 and 4, we wanted to test the differences between approach-avoidance training and approach-avoidance instructions paradigms. In the two experiments, participants were shown logos which were later associated with approach or avoidance. The association could be due to an approach-avoidance training or to mere approach-avoidance instructions. The logos were then used as primes during an AMP (Payne et al., 2005). As assumed by the APE model and single-process propositional models (but not the SEM), the primes associated with approach rather than with avoidance increased the likelihood for a target to be considered pleasant during the AMP, even when participants merely received approach-avoidance instructions (Experiment 3 & 4). This effect, however, did not differ significantly between the approach-avoidance training condition and approach-avoidance instructions condition (Experiments 3 & 4).

A notable contribution of these two experiments is the extension of results regarding the effect of approach-avoidance to a new procedure. The effects of both approach-avoidance training and approach-avoidance instructions on IAT performance abound in the literature (e.g., Kawakami et al., 2007, 2008; Phills et al., 2011; Van Dessel, De Houwer, Gast, et al., 2016; Van Dessel et al., 2015; Van Dessel, Gawronski, et al., 2017). To the best of our knowledge, however, we are the first to report effects of approach-avoidance training and approach-avoidance instructions on an AMP. This addition to the literature is important for several reasons. First, it is interesting regarding the automaticity features of the evaluative responses acquired because of approach-avoidance procedures. Automaticity is a complex construct (Moors & De Houwer, 2006) and several features of automaticity are involved when participants must take an indirect evaluative measure. Compared to an IAT procedure, the AMP does not require participants to process features of the evaluated stimuli, suggesting that the indirect evaluative responses resulting from approach-avoidance procedures could be

reactivated in a goal-independent fashion. Second, the sequential priming structure of the AMP offers opportunity to answer new research questions. Indeed, because of the way IAT scores are computed, they are doomed to be relative indices (Blanton & Jaccard, 2006). The same is not true for sequential-priming-structured paradigm. As long as neutral stimuli are included as prime in the procedure (Wentura & Degner, 2010), researchers can disentangle the respective contributions of the different stimuli, for example the ones associated with approach and those associated with avoidance (but for an example of such procedure using the SC-IAT, see Van Dessel, Gawronski, et al., 2017).

Regarding the question of the contribution of the repeated experience that we want to address in the present dissertation, Experiments 3 and 4 were inconclusive. The two experiments revealed no significant difference. We must also note that the statistical procedures do not allow us to conclude on whether the contribution of repeated experience would have been negligible (Lakens et al., 2018; Wellek, 2010). Still, from a theoretical perspective, one of the results we observed is interesting.

When reviewing the theoretical models of evaluation, we presented a model more radical than the other regarding the question of the acquisition of indirect evaluative response. The SEM assumes that the acquisition of indirect evaluative responses such as a response pattern in an AMP is necessarily the result of associative processes (McConnell & Rydell, 2014; Rydell & McConnell, 2006). The fact that participants evaluated differently stimuli associated with approach than stimuli associated with avoidance as a result of mere approach-avoidance instructions falsify this prediction. While the acquisition of indirect evaluative responses because of approach-avoidance instruction is not a new result (e.g., Van Dessel et al., 2015), Experiments 3 and 4 at least suggest that the results obtained by Van Dessel et al. are not the result of an artifact caused by the procedures they used (i.e., an IAT).

Another observation has to do with one of the models that we reviewed. The single-process propositional models do predict differences between approach-avoidance training and approach-avoidance instruction, and it predicts that these differences should be larger in some situations. In their single-process propositional model, Van Dessel et al. (2018a) suggest that repetition in approach-avoidance training should increase the likelihood for participants to infer a link between a movement and the stimulus. A corollary of that assumption is that we should not observe a difference between approach-avoidance training and approach-avoidance instructions procedure if the impact on the evaluative response that we observe depends on a proposition that is very easy to infer. If we assume that the critical propositions were easy to infer in Experiment 4, our results would be coherent with the single-process propositional models. This also has implication for further study. In a nutshell, the difference between approach-avoidance procedures should be easier to observe in situations where it is less easy for participants to make the inference critical for evaluation acquisition.

Conclusion

In this chapter, we wanted to assess the difference between approach-avoidance training and approach-avoidance regarding indirect evaluative responses acquisition. Across two experiments, we showed that the two procedures were effective at forming evaluative response on an AMP (Payne et al., 2005), a result which was not predicted by the SEM (McConnell & Rydell, 2014). The test of the difference between the two procedures, however, has proven to be inconclusive.

Chapter 8

Approach-Avoidance Effects in Incidental Designs

The single-process propositional models of evaluation predict several situations in which the difference between approach-avoidance training and approach-avoidance instructions will be larger. When the link between a stimulus and a movement is not obvious, for example, it will be harder for participants to infer the critical proposition which is a necessary condition for a new evaluative response to emerge (Van Dessel et al., 2018a). In such a scenario, the contribution of the repeated pairing between a movement and a stimulus should be larger. In this chapter, we decided to focus on one of these situations where the relevant propositions are harder to acquire. Specifically, we decided to focus on incidental approach-avoidance training, that is, procedures where participants do not receive instructions about the target of the training. We will investigate whether the contribution of repeated approach and avoidance movement is larger in such situation compared to when participants do receive instructions related to the target of the training.

Approach-Avoidance Training Design

The applications of approach-avoidance training are numerous. A wide range of topics has been covered by interventions adopting an approach-avoidance training: math (Kawakami et al., 2008), intergroup relationships (Kawakami et al., 2007; Phills et al., 2011), alcohol use (A. M. Hahn et al., 2019), eating behaviors (Ferentzi et al., 2018), sexual behaviors (A. M. Hahn et al., 2019), gambling (Boffo et al., 2017), depression (Becker et al., 2019). Interestingly, the design of the approach-avoidance training procedures was not always the same across the studies. Some of them used what we will call an incidental design, and some of them used what we will call an instrumental design. To understand the

difference between the two, one must keep in mind that an approach-avoidance training is designed to target a behavior associated with a class of stimuli. For example, Boffo et al. (2017) designed a procedure designed to change behaviors related to gambling stimuli. Kawakami et al. (2008) designed a procedure to change behaviors related to math. The incidental or instrumental nature of the training indicates whether the approach-avoidance instructions refer to the stimuli targeted by the intervention.

For instance, in Boffo et al.'s (2017) study, participants did not have to approach and avoid gambling related stimuli. Rather, they were asked to approach pictures tilted to the left and avoid pictures tilted to the right. Without controlling how many gambling related pictures were tilted to the left or to the right, participants could have ended up approaching gambling related pictures, avoiding gambling related pictures, or both. This was not true for Kawakami et al.'s (2008) participants who received the specific instruction to approach (avoid) math-related pictures. Of course, because Boffo et al.'s (2017) study was designed to target gambling related behavior, participants did avoid gambling related stimuli. Because the link between a class of stimulus and a movement is given to participants in the instructions in instrumental design, it should be easier for participants to acquire the relevant propositions when the design of the experiment is instrumental—such as the one used by Kawakami et al. (2008)—rather than when it is incidental—such as the one used by Boffo et al. (2017).

Incidental Approach-Avoidance Training

One of the first lines of research to use an incidental design was conducted by Woud et al. (2008, 2011, 2013). In their papers, the researchers were interested in the learning of evaluative responses toward faces. Specifically, they wondered whether incidental approach-avoidance training of faces resulted in the acquisition of an evaluative response. In other words, does approaching a face makes us evaluate it more positively, even when we are not explicitly instructed to approach that face. In the paradigm Woud et al. (2008, 2011, 2013)

designed, participants were instructed to approach and avoid images according to the color of their background (e.g., “approach pictures that have a blue background and avoid pictures that have a brown background”). In doing so, by design, participants always approached the same faces (e.g., faces A and B) and avoid the same faces (e.g., faces C and D). Then, participants performed different evaluative tasks such as evaluative priming (Woud et al., 2008, 2011, 2013) or explicit face rating (Woud et al., 2008). Woud et al.’s (2008, 2011, 2013) results revealed that the faces which had been approached were evaluated more positively than the faces which had been avoided. In addition to providing support for an effect of incidental approach-avoidance training, this line of research also highlighted several moderators of this effect, notably the number of pairings between a face and a movement (the more trials in the approach-avoidance training, the stronger the acquired evaluative responses; Woud et al., 2011).

The results of Woud et al. (2011) regarding the effect of repetition during approach-avoidance training aligns with Van Dessel et al.’s (2018a) predictions. In their single-process propositional model, Van Dessel et al. (2018a) describe proposition formation as an attention-consuming process. With repetition, the likelihood that a critical proposition is acquired increases. This effect, however, should not be observed for propositions particularly easy to acquire—we could expect a ceiling effect of the number of repetitions. Woud et al.’s (2011) paradigm, compared to the one offered by Boffo et al. (2017) as remediation to gambling addiction, represents a unique opportunity to investigate the effect of repeated experience in approach-avoidance paradigm. Indeed, with Woud et al.’s (2011) paradigm, participants have to form one proposition per faces for the training to be effective. In Boffo et al. (2017), they have to form one proposition per category of stimulus. The complexity of acquisition of the critical propositions should correlate with the sensitivity to detect an effect of pairing

repetition. We decided to make a few adjustments to this paradigm to investigate the effect of repeated experience on approach-avoidance training.

Improving Woud et al.'s Paradigm

In order to investigate the effects of incidental approach-avoidance training procedures, as in our previous experiments, we decided to use an AMP rather than an evaluative priming task (Payne et al., 2005). Indirect measures—such as the AMP, the IAT, the evaluative priming task, or the go/no-go association task—differ in terms of reliability (Gawronski & Hahn, 2019). The proportion of noise (rather than signal) captured by the measure can be higher according to the measure that is used. In their experiments, Woud et al. (2008, 2011, 2013) investigated the effects of the incidental approach-avoidance training with an evaluative priming task (Fazio et al., 1986), a measure known for its low reliability (Gawronski & Hahn, 2019). Because of reports challenging the robustness of Woud et al.'s (2008) findings (Vandenbosch & De Houwer, 2011), we decided to use an AMP instead of an evaluative priming task. This procedure shows a high reliability (Payne & Lundberg, 2014) and we know that it can capture the effects of approach-avoidance procedures (Experiments 3 & 4).

Another advantage of using the AMP is its ability to capture variance at a stimulus level (Wolsiefer et al., 2017). Statistical analysis of indirect evaluative responses allows researchers to assess what is causing variation in a stimulus evaluation. Using mixed model analyses, Wolsiefer et al. (2017) compared the source of variance in evaluative priming task (Fazio et al., 1986), the IAT (Greenwald et al., 1998) and the AMP (Payne et al., 2005), revealing that the AMP is better suited to capture the variance which is related to stimuli. In other words, the differences between the different primes make a bigger impact on responses in an AMP compared to an evaluative priming task. Because we are interested in testing

whether participants can generate the propositions necessary for an evaluative response to emerge for each of the stimuli, this feature of the AMP appeared critical.

To make an experiment investigating the effects of approach-avoidance procedures in an incidental design as comparable as we could to our previous experiments (Experiments 1a and 1b), we decided to use words as our stimuli of interest. In their experiments, Woud et al.'s (2008, 2011, 2013) had to approach and avoided faces according to the color of the image background. In our experiment, we decided to make participants approach and avoid words according to their color: Participants will be instructed to approach words in pink and avoid words in orange or the opposite.

Experiment 5

To sum up, in this experiment, we wanted to assess the contribution of repeated experience when evaluations were relying on a hard-to-acquire proposition. We believe that such a situation should increase the contribution or repeated experience in evaluative learning. To assess the contribution of repeated exposure, as in Experiment 3, we will rely on two groups of participants. In a first part of the experiment, participants will be randomly assigned to an approach-avoidance training condition and in an approach-avoidance instructions condition. They will be instructed to approach or avoid words based on their colors. After the approach-avoidance procedures, in a second part of the experiment, participant will take an AMP. In this AMP, we decided to include stimuli whose evaluation was dependent on hard-to-make inference, but also stimuli whose evaluation was dependent on easy-to-make inference (see Table 8.1). Stimuli whose evaluation depended on hard-to-make inferences will be similar to the stimuli in an incidental approach-avoidance training. That is, participants will not receive approach-avoidance instructions targeting specific words, but they will end up approaching and avoiding them during the training. Stimuli whose evaluation depended on easy-to-make inferences are similar to stimuli in an

instrumental approach-avoidance training. The approach-avoidance instructions participants will receive will target those stimuli.

Table 8.1

Characteristics of the Words used as Primes during the AMP

Type of features of the stimulus	Evaluation depends on propositions	Color of the stimulus	Words	Implementation in the experiment
Instrumental in the training	Easy to acquire	Associated to approach-avoidance instructions	Not approached and avoided during the training	New word of same color
Incidental to the training	Hard to acquire	Not associated to approach-avoidance instructions	Approached and avoided during the training	Same words of new color

We expected the evaluative responses acquired as a result of approach-avoidance training to be of a larger magnitude than those acquired as a result of approach-avoidance instructions. We expected this difference between the two procedures to be larger when the stimuli had a feature incidental to the training (i.e., when the evaluation depended on hard to acquire propositions) rather than when the stimuli had features instrumental in the training (i.e., when the evaluation depended on easy to acquire evaluation). This experiment was preregistered. We report all measures, manipulations, and exclusions. Material, data, and analysis scripts can be found at <https://osf.io/5tcn4/>.²⁷

In this experiment, we investigated the contribution of repeated experience to incidental evaluative learning. To do so, our participants received the instruction to approach

²⁷ This repository might be subject to an embargo. If it is unavailable, please e-mail cedric.batailler@gmail.com.

and avoid two lists of words based on their color. To assess the contribution of repeated experience, only half of the participants completed the approach-avoidance training. The other was assigned to an approach-avoidance instruction condition. We then measured evaluative response with an AMP for two kinds of stimuli. The first kinds of stimuli possessed features which were instrumental for the training, that is, participants specifically received instructions to approach or avoid them. The second kinds of stimuli possessed features which were incidental to the training, that is, participants in the training condition ended approaching and avoiding them, even if they did not receive the instructions to do so—these stimuli were colorless versions of the words which had been approached and avoided during the training. These stimuli were also evaluated by participants in the approach-avoidance conditions to serve as a comparison baseline.

Method

Participants, Sensitivity Analysis, and Design

Two hundred and thirty-eight ($M_{age} = 20.83$, $SD_{age} = 2.88$, 209 women, 29 men) psychology students from the Université Grenoble Alpes (France) ended up taking part in this study in exchange for course credit.²⁸ We relied on a 2-condition design (approach and avoidance instructions vs. approach and avoidance training).

We pre-registered a sample size of 200 participants for this experiment. With this sample size, we had a statistical power of 29% to detect small between-subject effect sizes ($\eta^2_p = .010$) a statistical power of 94% to detect medium between-subject effect sizes ($\eta^2_p = .060$), and a statistical power above 99% to detect large between-subject effect sizes

²⁸ We actually collected 239 unique participations to our study, but one participant completed the experiment twice. We decided to exclude their second participation from our analysis.

($\eta^2_p = .140$; Cohen, 1988). The decision to adopt this sample size was based on resource constraints and feasibility.

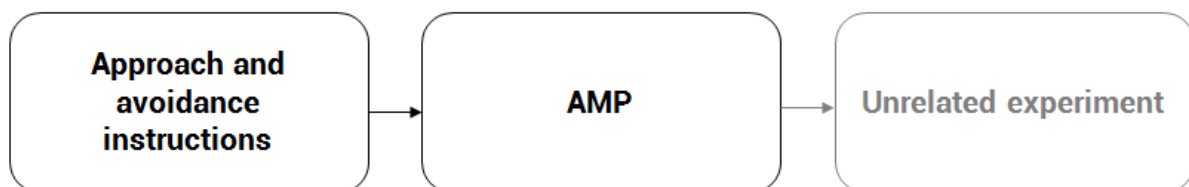
Procedure and Material

One out of two experimenters welcomed the participants in the lab, in a room where up to eight participants could take the study (participants were separated from each other by a divider). This experiment was entirely computerized and was coded using E-Prime. Participants were randomly assigned to either the approach-avoidance training or the approach-avoidance instructions condition. The experiment was included in a battery for participants in the approach-avoidance instructions condition, as it took participants less time to complete it. Experiment 5 was always the first experiment of the battery (see Figure 8.1). Experimenters were blind to the experimental condition of participants.

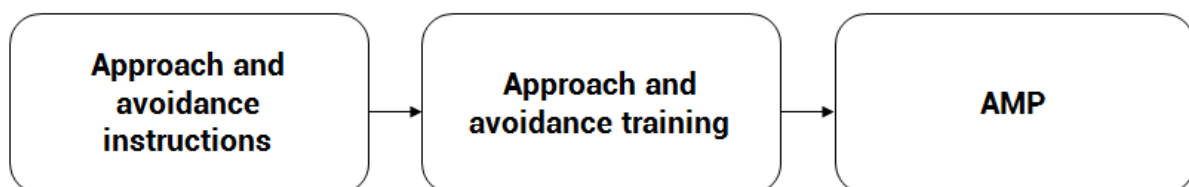
Figure 8.1

Time Course of Experiment 5 According to the Experimental Condition

Approach and avoidance instructions condition



Approach and avoidance training condition



As in our other experiments, participants first received approach-avoidance instructions. In this experiment, they would see two different lists of words: the orange list

and the pink list. These lists were composed of 5 non-words randomly selected from a list of 20. Participants had to approach words from the orange list and avoid the words from the pink one or the opposite (this instruction was counterbalanced across participants). Each word of the colored lists was presented during the instructions (see Table 8.2). As with any of our experiments including an approach-avoidance instruction condition, participants were told that they would learn how to perform approach and avoidance movement later. For now, they had to remember which color went along with which movement because this information was critical for the completion of the study. After receiving these instructions, participants in the approach avoidance training condition completed the training while participants in the instructions condition completed the AMP.

As in Experiments 3 and 4, participants in the approach-avoidance training condition completed the VAAST-based training which we developed in **Chapter 7**. Participants approached and avoided stimuli within a virtual environment on the computer and they used the keyboard arrow keys of the directional keypad to do so. Each trial started with a fixation cross and was followed by a colored target word. Participants had to press the correct response key twice to complete a trial, each key press triggering an approach or avoidance animation. Errors resulted in feedback inviting the participants to complete the trial again. Participants started with a 20-trial practice block and then moved to a 200-trial main block. The approach/avoidance phase lasted approximately 10 minutes. Participants in the training condition performed the AMP after this training.

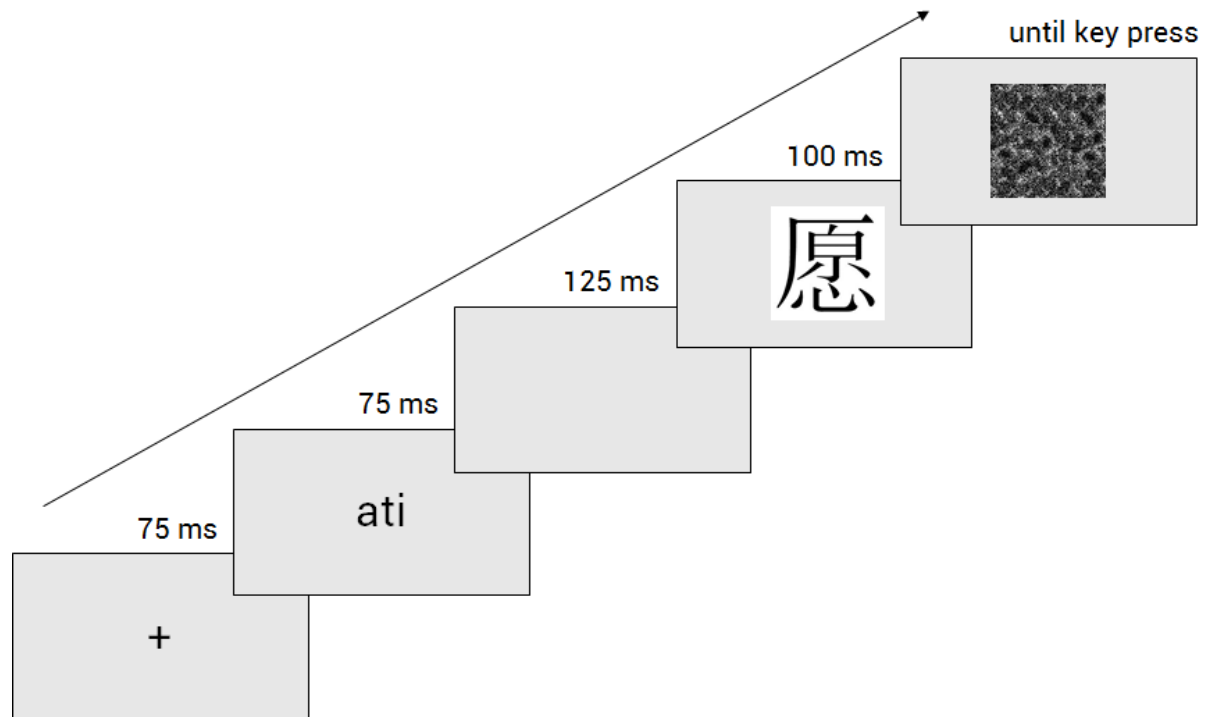
Both participants in the approach-avoidance training and approach-avoidance instructions condition completed an AMP (Payne et al., 2005). As in Experiments 3 and 4, this task was presented as a way to understand how individuals do simple but quick judgments. In every trial, they had to evaluate whether a Chinese ideogram (the target) was pleasant compared to an average one. Critically for the present experiment, this target was

preceded by a prime which was associated with either approach or avoidance (see Figure 8.2). Critically for our hypotheses, two kinds of features could associate a stimulus to a movement. The first kind of stimulus was a novel word in either pink or orange color. We consider that these stimuli possess features which are *instrumental* for the training. If these stimuli had been presented during the training participants should have approached or avoided them. The second type of stimuli possessed features which were not instrumental for the training instructions. These stimuli were the same words as the one that were actually approached or avoided (or for which participants actually received approach-avoidance instructions), but they were presented in black. As participants did not receive instructions related to black stimuli but received instructions to approach and avoid those words, we consider these stimuli as possessing feature which were *incidental* for the training (see Table 8.2). To sum up, the stimuli which had the same color were the stimuli with instrumental features and the stimuli which were the same words were the stimuli with incidental features. These two types of stimuli were randomly selected until participants completed 72 trials. Reliability of the measure was computed using the same calculation as in Payne et al. (2005) and reached $\alpha = .59$, 95% CI [.51; .67].²⁹

²⁹ Reliability of the AMP when stimuli possessed features instrumental for the training was $\alpha = .71$, 95% CI [.65; .76]. Reliability of the AMP when stimuli possessed features incidental to the training was $\alpha = .18$, 95% CI [.03; .34].

Figure 8.2

Time Course for a Trial of the AMP in Experiment 5



Note. In this trial, the prime stimulus possesses features that are incidental for the approach-avoidance training. Participants had to say whether they consider the Chinese ideogram pleasant or not, compared to an average Chinese ideogram.

Table 8.2

Example of Stimuli Used in Experiment 5

Training and instructions	AMP — Incidental features	AMP — Instrumental features
ati (orange)	ati (black)	—
beto (orange)	beto (black)	—
cela (orange)	cela (black)	—
duvi (orange)	duvi (black)	—
laso (orange)	laso (black)	—
ceno (pink)	ceno (black)	—
fiza (pink)	fiza (black)	—
lebu (pink)	lebu (black)	—
nole (pink)	nole (black)	—
nore (pink)	nore (black)	—
—	—	eska (orange)
—	—	jiri (orange)
—	—	ome (orange)
—	—	tope (orange)
—	—	wena (orange)
—	—	ivo (pink)
—	—	miti (pink)
—	—	nenno (pink)
—	—	pita (pink)
—	—	zalu (pink)

Note. Parentheses indicate the color in which the stimulus was displayed.

At the end of the experiment, participants were asked for demographic information. They were then thanked and debriefed by the experimenter. In total, the experiment lasted 25 minutes.

Results

In this experiment, we were interested in the effect of the movement associated to a prime on the evaluation of an AMP target. Precisely, we wanted to test whether the contribution of repeated experience was larger for stimuli with features incidental to the training than for stimuli with features instrumental for the training. If so, the difference between the approach-avoidance training and approach-avoidance instructions procedures should be larger when the primes possess features incidental to the task, compared to when they possess features instrumental.

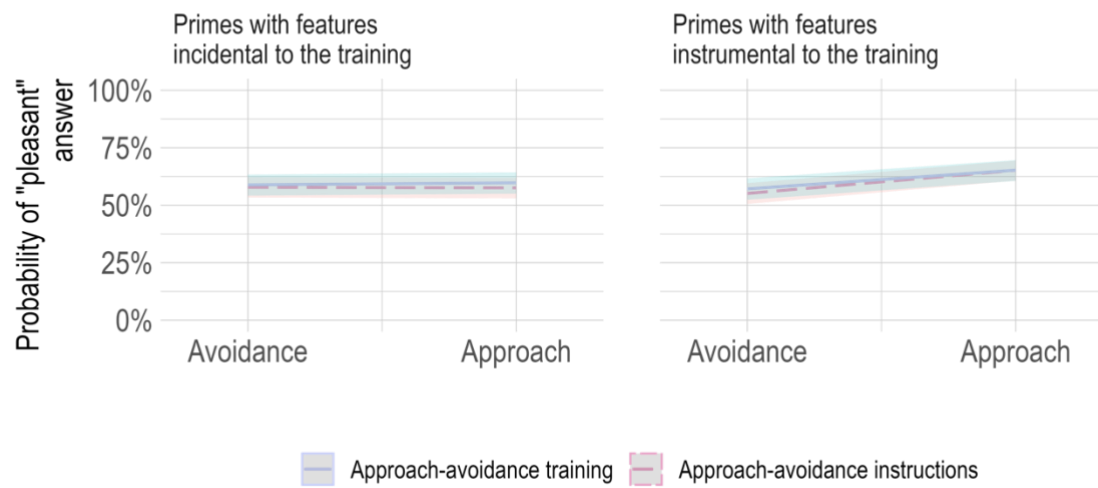
As preregistered, we used *lme4* to conduct a binomial mixed model (Bates, Mächler, et al., 2015). In this model, we were interested in the factors making participants more likely to answer “pleasant” than “unpleasant” during the AMP to our data (Bates, Mächler, et al., 2015). Fixed effects predictors were the movement associated with the stimulus (approach vs. avoidance), whether the participants completed the training or not, and the type of features the stimulus possessed (instrumental for the training vs. incidental to the training). We used participants and the stimuli as random effects for this analysis. To deal with convergence issues, we followed the recommendation by Bates et al. (2015) to simplify the model (see Appendix 4). Data from the 238 participants responding (several times) to the 20 stimuli made a total of 16,184 observations to fit the model. For readability reasons, we will only report effects which are theoretically relevant in the main text. Exhaustive results can be found in Table 8.3.

The mixed model analysis first revealed that participants were 1.48 times more likely report to answer “pleasant” than “unpleasant”, $z = 8.60$, $p < .001$, 95% CI [1.35; 1.62]. More importantly, the main effect of movement was also significant, participants were 1.22 times more likely to consider the targets as pleasant when they were preceded by a prime associated with approach rather than avoidance, $z = 3.19$, $p = .001$, 95% CI [1.08; 1.38]. This main effect was qualified by the features the stimuli possessed. The effect of prime’s movement was significantly stronger when the stimulus possessed features instrumental for the training rather than features incidental for the training, $z = 3.47$, $p < .001$, $OR = 1.54$, 95% CI [1.17; 1.78]. Simple slope analysis revealed that, when the stimulus possessed features instrumental for the training, participants were 1.47 times more likely to consider the target answer pleasant when the stimulus was associated with approach rather than avoidance, $z = 3.86$, $p < .001$, 95% CI [1.21; 1.78]; the effect of the movement associated with the prime was not significant when stimulus possessed features incidental to the training, $z = 1.01$, $p = .825$, $OR = 1.01$, 95% CI [0.90; 1.14] .

Finally, the critical interaction between the movement associated with the primes, the prime features, and the type of training participants received was not significant. The difference of movement effect between stimuli possessing instrumental and incidental features was not significantly different between participants who had to perform the approach and avoidance training and those who received instructions only, $z = -0.60$, $p = .550$, $OR = 0.88$, 95% CI [0.58; 1.34]. In other words, participants are not significantly more likely to develop evaluative responses for stimuli possessing features incidental to the training when they actually performed the training compared to when they received instructions only (see Figure 8.3).

Figure 8.3

Effects of the Primes During the AMP According to their Features and the Experimental Condition of Participants



Note. Confidence interval represents the 95%CI of prediction at the population level based on the mixed model.

Table 8.3*Summary Statistics of Experiment 5's Mixed Model*

	log(OR)	95% CI	z	p-value
(Intercept)	0.39	0.30, 0.48	8.26	< .001
Movement	0.2	0.08, 0.32	3.19	0.001
Training	0.05	-0.13, 0.24	0.57	0.567
Features	0.09	-0.02, 0.21	1.57	0.118
Movement × Training	-0.01	-0.23, 0.20	- 0.14	0.89
Movement × Features	0.37	0.16, 0.58	3.47	< .001
Training × Features	-0.03	-0.26, 0.21	-0.21	0.833
Movement × Training × Features	-0.13	-0.54, 0.29	-0.60	0.551

Note. OR = Odds ratio. CI = Confidence Interval.

As pre-registered, we also conducted a by-participant analysis. Because this analysis is convergent with the mixed-model one, in terms of direction of the effect as well as in terms of significance, we will not report it, but it can be found online on the OSF repository of this experiment (<https://osf.io/5tcn4/>).

General Discussion

In this chapter, we wanted to use an approach-avoidance training with an incidental design to investigate the contribution of repeated experience. In this experiment, 238 participants completed an AMP capturing evaluative responses related to two kinds of features of stimuli. The first kind of feature was instrumental for the approach-avoidance training—participants received specific instructions to perform a movement when confronted to this feature, the second kind of feature was incidental to the training—participants in the

training condition did not receive any instructions regarding this feature but ended up approaching and avoiding stimuli who possessed them. Our two groups of participants, which either completed the training or merely received approach-avoidance instructions, allowed us to estimate the contribution of repeated experience to the acquisition of evaluative responses toward stimuli who possessed features which were either instrumental or incidental to the training. While our experiment revealed that approach-avoidance procedures resulted in the acquisition of an evaluative response toward stimuli which possessed features instrumental for the training, we found no evidence that it impacted the evaluation of stimuli which possessed features incidental to the training. We did not find evidence either that repeated experience over mere approach-avoidance instructions contributed to evaluative response acquisition.

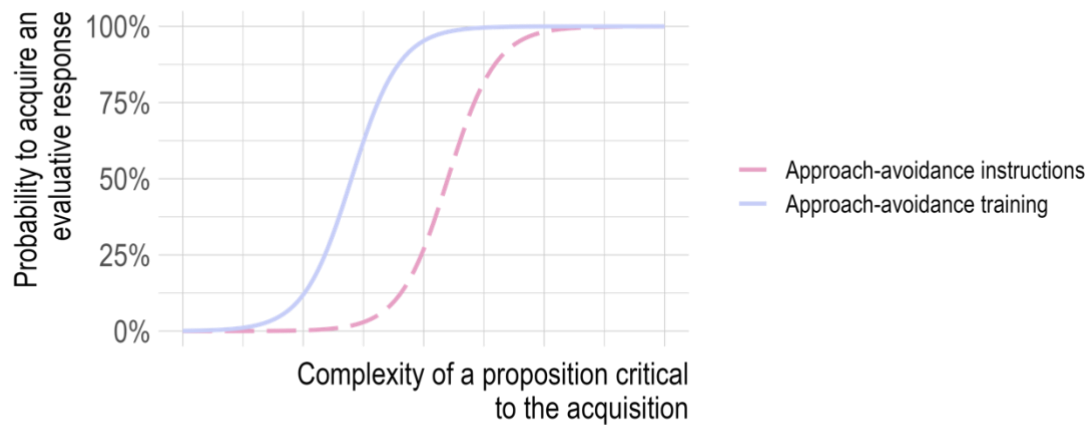
In Experiment 5, our participants acquired indirect evaluative responses for stimuli which possessed features instrumental for the training, but not for stimuli which possessed features incidental for the training. One might wonder whether our results questions procedures which used incidental approach-avoidance training (e.g., Becker et al., 2019; Boffo et al., 2017; Ferentzi et al., 2018; A. M. Hahn et al., 2019). The experiment we conducted differs from the approach-avoidance training used with clinical application in several ways. We can easily assume that the acquisition of several propositions is a more resource-consuming processes than the acquisition of one proposition. In Experiment 5, if we adopt the perspective of a single-process propositional model, participants must acquire one proposition per stimulus for an indirect evaluative response to emerge (with our material, 10 in total). This is different from the incidental approach-avoidance trainings which are used with clinical application and which target a whole class of stimuli. These trainings require that participants acquire one proposition for an indirect evaluative response to emerge (e.g., “I am avoiding gambling-related stimuli”, Boffo et al., 2017). Our pattern of results is

consistent with this idea, as our participants acquired evaluative responses for stimuli possessing features instrumental for the training, and they did not for stimuli possessing features incidental for the training. This would align with the literature reviews highlighting the non-efficient nature of indirect evaluative learning (Corneille & Stahl, 2019).

In Experiment 5, we decided to adopt an incidental design with the idea that increasing the complexity of the acquisition of a critical proposition would also increase the sensitivity to detect the contribution of repeated experience. In the discussion, we suggest that increasing the complexity of a proposition critical to the acquisition of an indirect evaluative response might have been too complex for participants. At first, these assumptions could appear contradictory, however they are not. If we assume that the complexity of a proposition is associated with the probability that a participant acquires an evaluative response, and if repeated experience makes the acquisition of the evaluative response easier, both ceiling and floor effects hiding the effect of repeated experience can emerge (see Figure 8.4). This model is empirically testable, it would be consistent with a quadratic relationship between the critical proposition complexity and the difference between approach-avoidance training and approach-avoidance instructions in terms of evaluative response acquisition.

Figure 8.4

Relationship Between the Complexity of Acquisition of a Proposition and the Acquisition of an Evaluative Response



Note. Propositions on the right are less complex than those on the left. The sensitivity to detect a difference between approach-avoidance training and approach-avoidance instructions is associated with the vertical difference between the two curves. Solid curve represents approach-avoidance training and dashed one represents approach-avoidance instructions.

From a theoretical perspective, it is interesting to note that, as in Experiments 3 and 4, our results are hard to reconcile with the predictions of the SEM (McConnell & Rydell, 2014; Rydell & McConnell, 2006). If anything, this model would have predicted an effect of repeated experience, and no effect of mere instructions. This prediction does not fit with the effect of approach-avoidance instructions on the acquisition of evaluative responses for stimuli which possess features instrumental for the training. Regarding the APE model (Gawronski & Bodenhausen, 2006, 2007, 2014a), a critical point is the acquisition of an indirect evaluative response as a result of repeated experience. While consistent with our results regarding stimuli with features instrumental to the training, the APE model also

predicted an effect for stimuli with features incidental to the learning, an effect we did not observe. The statistical procedures that we used, however, do not allow us to conclude on the basis of a null effect (Lakens et al., 2018). In the end, it seems that the pattern of data obtained in Experiment 5 is more easily explained by single-process propositional models of evaluation than by dual models.

Conclusion

In this chapter, we wanted to assess the effect of repeated experience in approach-avoidance training by using a procedure with an incidental design. We thought that the adoption of this design would increase the chance that we observe an effect of repeated experience. We could not, however, find evidence that any of the approach-avoidance procedure used yielded to the acquisition of an evaluative response toward stimuli whose features were not mentioned in the approach-avoidance instructions. Results regarding evaluative learning regarding stimuli possessing features instrumental for the training are consistent with **Chapter 7**—both approach-avoidance training and approach-avoidance instructions result in the acquisition of indirect evaluative response.

Chapter 9

Effects of Approach-Avoidance Procedures Over Time

As discussed in the introduction section, early views on indirect evaluative responses considered them as the results of our many interactions with the world, recorded and now deeply rooted in memory (Fazio et al., 1986; Greenwald & Banaji, 1995). In this dissertation, we use approach-avoidance training procedures as a way to investigate the contribution of repeated experience in the acquisition of indirect evaluative learning. In each of our experiments (Experiments 1a–5), so far, participants took an indirect evaluative response task right after completing the approach-avoidance procedures. We have not yet examined the effect of the approach-avoidance procedure on a longer period of time. In that sense, these experiments—whose design is common in the implicit social cognition literature (Forscher et al., 2019)—differ a bit from the early implicit social cognition work where indirect evaluations were thought to be influenced by past experience rather than by what just happened. In the present chapter, we decided to focus on the lasting effects of approach-avoidance training. Specifically, we decided to test whether indirect evaluative responses acquired as a result of an approach-avoidance training were retrieved the day after, and whether the repeated experience of approach and avoidance during a training contributed to this sustainability of acquired evaluative responses. As we will see, investigating the effects of approach and avoidance training in a longer term offers interesting perspective in the investigation of the role of repeated experience.

On the Stability of Indirect Evaluative Responses

In the introduction section of this dissertation, we discussed how indirect evaluative responses such as the evaluative priming task or the IAT appeared in the arsenal of social

psychologists. To overcome some of the limits with self-report measures, some researchers used the progress made on automaticity and implicit memory to design measures with the goal to capture something deeply rooted in our minds. For Greenwald and Banaji (1995), deeply rooted in the mind meant that the traces from the past (which are no longer identifiable) now influence our behaviors; For Fazio et al. (1986), deeply rooted meant that what was encoded in the past re-emerges automatically. Consistently, at the time, what was captured by indirect evaluative responses was considered as remarkably stable (Devine, 1989). When focusing on intergroup relationship, what was captured by indirect evaluative response—and was influencing prejudice behavior—was even considered as a “Cognitive Monster” that it was impossible to get rid of (Bargh, 1999). These views were coherent with the idea that indirect evaluative responses captured something which was progressively encoded in long-term memory over time and integrated in our behaviors.

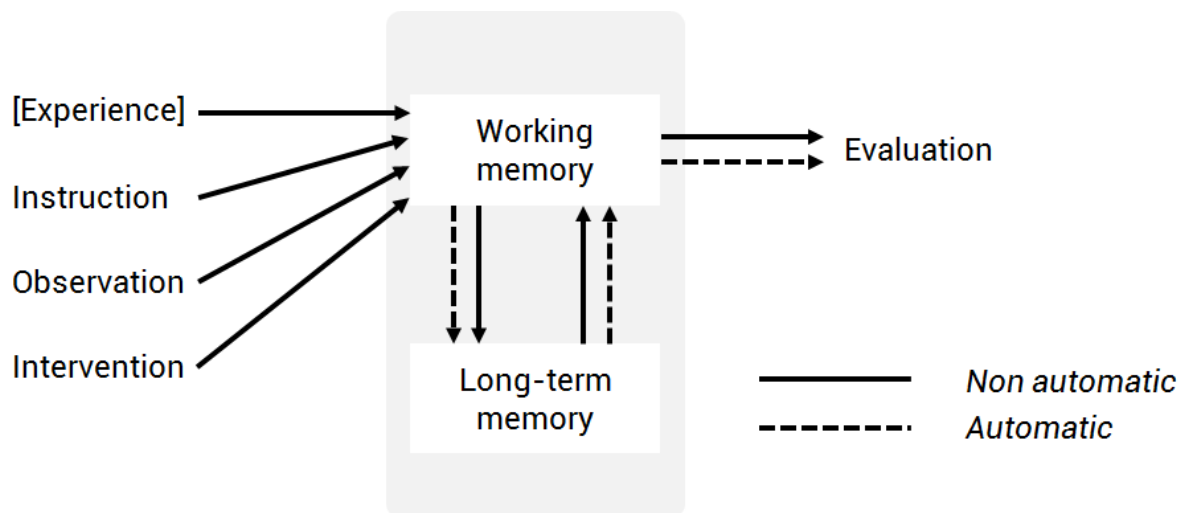
Since then, however, the advances in implicit social cognition changed the views of researchers on the question of the stability of indirect evaluative responses. Specifically, they are now considered as more malleable than initially thought (see Blair, 2002). Among the many examples showing that someone’s indirect evaluative response could change, Lowery et al. (2001) showed that scores at a racial IAT (capturing association between the Black-White and Good-Bad concepts) depended on whether the experimenter was a Black or White American, Kawakami et al. (2007) that an approach training of Black American impacted the same racial IAT, and M. J. Ferguson (2008) that subtly activating a goal could change evaluative responses in a goal-consistent manner. Nowadays, models of evaluations such as the single-process propositional model account for the data related to the stability and the data related to the malleability of indirect evaluative responses (De Houwer, 2018).

Long-Term Encoding and Indirect Evaluative Responses

Single-process propositional model of evaluation consider propositions as the basis of information. These propositions—that is, mental representations qualifying the relationship between concepts (see the Propositional Processes section)—emerged in working memory, based on repeated experience, instructions, or other external events which are analyzed and combined to form the propositions (see Figure 9.1; De Houwer, 2018). Once a proposition is processed in working memory, it can influence evaluation. Yet, the storage capacity of working memory is limited (Miller, 1956), and propositions need not live permanently in working memory for them to influence evaluation. Some of them will be encoded in long-term memory and retrieved at the appropriate time.

Figure 9.1

Single Process Propositional Model of Evaluative Learning



Note. Adapted from De Houwer (2018).

Propositions are encoded for long term retrieval in what is called memory traces (e.g., Hintzman, 1984; Versace et al., 2014). Memory traces contain episodic information about a

situation such as the perceptions an individual has, or the mood they are in, or, according to the single-process propositional model of evaluation, “information about the ways in which stimuli and responses are related in that situation” (De Houwer, 2018, p. 9). Once encoded in long-term memory, the traces will be later reactivated on the basis of similarity between the information contained in the trace and a current situation. Once reactivated, De Houwer (2018) suggests that the traces will be available again in working memory. Because episodic traces are assumed to store propositions, and because they can be reactivated, propositions can influence evaluative responses in the long run.

This model can account for both the stability and malleability of the indirect evaluations. If a memory system encodes many memory traces containing consistent propositional information, the evaluations produced by an individual should be stable over time. This is true as long as the propositions encoded in long term memory emerge in working memory during the evaluation. If, however, these propositions do not find their way to working memory, then the evaluative response emitted should differ from the usual. We could make the hypothesis that an approach-avoidance training influences the propositions considered in working memory (Van Dessel et al., 2018a). If, right after the training, someone performs an evaluative measure which depends on the content of working memory, it should be influenced by the training. However, if someone performs the measure after a delay, the influence the training could have depended on whether the relevant propositions have been encoded in long-term memory. A change in evaluative response could be momentary or not.

Approach-Avoidance Training to Sustainably Acquire Evaluative Responses

In our previous experiments (Experiments 1a–5), each time participants engaged with an evaluative response measure, the measure was included right after the approach-avoidance procedure (be it an approach-avoidance training or approach-avoidance instructions).

Because of this temporal proximity between the two events, we can make the assumption that working memory by itself can account for evaluative response acquisition (see figure 9.1, p. 177; De Houwer, 2018). In other words, a period that short between the approach-avoidance procedures and the indirect evaluative response measure does not require participants encoding the propositions in long-term memory for these interventions to have an effect. The same would not be true if the duration between the two events was larger or if participants had to use their working memory for an unrelated task in between them (Baddeley, 2010; Baddeley & Hitch, 1974).

In the present dissertation, we are focusing on the contribution of repeated experience in evaluative learning. To do so, we compared approach-avoidance training and approach-avoidance instruction procedures. If approach-avoidance training and approach-avoidance instruction procedures (only) differed regarding their ability to help the encoding of information in long-term memory, we would not have been able to observe difference between the two procedures (Experiments 3–5). Yet, the hypothesis of a difference between approach-avoidance training and approach-avoidance instruction procedures regarding their effects on long-term encoding seems reasonable.

In her Declarative Memory Model, Gast (2018) reviews the factors related to memory processes which play a role in evaluation acquisition (focusing on Evaluative Conditioning). Some of the factors related to the learning process (rather than the retention of information or its retrieval) could differ between approach-avoidance training and instruction procedures. For example, we can make the hypothesis that the attention allocated the target stimuli differs across the procedures. Because an approach-avoidance training, by design, requires participants actively interacting with the different targets, participants should show a higher level of attention during approach-avoidance training compared to approach-avoidance instruction procedures. The attention participants recruit to process the target stimuli should

help with the encoding of the relevant information in long-term memory (Fisk & Schneider, 1984). Another relevant factor could be the level of concreteness, which should be higher in approach-avoidance training, as movement are implemented more realistically. This level of concreteness should help with the encoding of the relevant information (Gast, 2018). For these reasons, we decided to compare the evaluative responses acquired as a result of approach-avoidance training and approach-avoidance instructions over time. In this experiment, we decided to focus on an identification IAT. Unlike Experiment 4 and 5, the design Experiment 6 does not require us to adopt a procedure with a structure similar to evaluative priming task. We decided to focus on an identification IAT other studies showing that performance on an identification IAT mediated the performance on other evaluative responses (Van Dessel, Gawronski, et al., 2017).

Experiment 6

In Experiment 6, we decided to investigate the effect of time on evaluative learning resulting from approach-avoidance procedures. The goal of this study was to investigate whether approach-avoidance training and approach-avoidance instructions produce different evaluative responses, after a delay. To do so, we used a paradigm in which participants first engaged with either an approach-avoidance training or approach-avoidance instructions procedure. Right after, we checked whether this procedure resulted in the acquisition of an evaluative response with an identification IAT. Participants took the same IAT the day after to see how the acquired evaluative responses evolved over time. We were particularly interested in whether evaluative responses acquired as a result of approach-avoidance instructions lasted as much as evaluative responses acquired as a result of an approach-

avoidance training. We report all measures, manipulations, and exclusions. Material, data, the analysis script, and the preregistration can be found online at <https://osf.io/etsmv/>.³⁰

Method

Participant, Design, and Sensitivity Analysis

Four hundred and four participants took part in this experiment in exchange of 3.05£ ($M_{\text{age}} = 36.6$, $SD_{\text{age}} = 13.1$; 230 female, 170 male, and 4 unreported gender). Participants were recruited using the Prolific platform (Palan & Schitter, 2018). They had to be native English speaker, to live in the UK, and had to have an approval rate over .90 on the Prolific platform to take part in our experiment. We recruited participants from the UK to make the study easier to run, as this would imply having to deal with only one time zone.

Participants engaged in a study with one between-subject variable (the type of interventions: approach avoidance training vs. approach-avoidance instructions) and one within-subject variable (the time of the IAT: the intervention day vs. the day after the intervention). The types of stimuli participants had to approach was counterbalanced across participants.

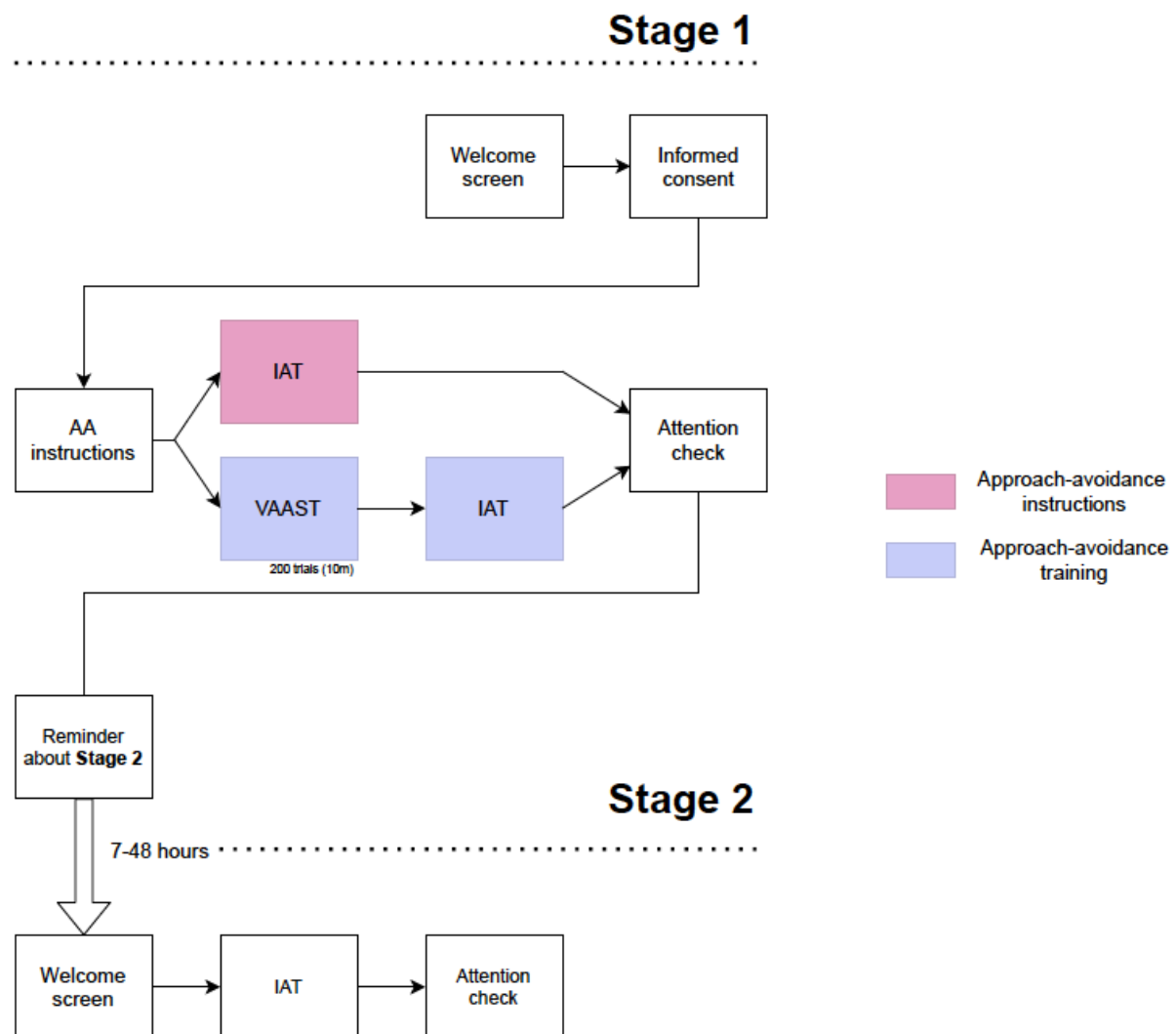
We preregistered a total sample size of 400 participants. Because we considered the risks of attrition higher for a two-days online experiments, our stopping rules imposed us to keep recruiting participants until the data from at least 400 participants were useable. With such a sample size, we had .51 statistical power to detect small between subject effects ($d_s = 0.20$), over .99 statistical power to detect medium between subject effects ($d_s = 0.50$), and over .99 statistical power to detect large between subject effects ($d_s = 0.80$; Cohen, 1988). Regarding within-subject effects, we had .97 statistical power to detect small effects

³⁰ This repository might be subject to an embargo. If it is unavailable, please e-mail cedric.batailler@gmail.com.

($d_z = 0.50$), over .99 statistical power to detect medium effects ($d_z = 0.50$), and over .99 statistical power to detect large effects ($d_z = 0.80$). The decision to adopt this sample size was based on resource constraints and feasibility.

Procedure and Material

We recruited our participants for a two-day experiment using the Prolific platform (Palan & Schitter, 2018). Participants knew that they would have to come back the day after to complete the experiment. The first stage of the experiment was made available at 7:00 GMT+1 and participant could sign up until 23:59 GMT+1. We will first describe this first stage of the experiment and then the second. A representation of the full course of the experiment can be found in Figure 9.2.

Figure 9.2*Course of Experiment 6*

Note. AA: Approach-avoidances, IAT: Implicit Association Test, VAAST: Visual Approach-Avoidance by the Self Task.

Stage 1. Upon the first connection to the webpage hosting Stage 1, participants were randomly assigned to an approach instruction condition (approach Luupite or approach Niffite) and to an approach-avoidance procedure condition (approach-avoidance training or approach-avoidance instructions). Accordingly, they were informed that the first stage of the

experiment would last 20 minutes (approach-avoidance training condition) or 7 minutes (approach-avoidance instruction condition). Participants were notified that this experiment might include attention checks. This experiment was coded using jsPsych (de Leeuw, 2015) and was automatically set on Fullscreen so that participants could focus entirely on the study.

The experiment was presented as a two-day experiment which included several tasks. Participants were informed that these tasks implied interacting with two sets of stimuli: the Luupite set and the Niffite set. A full list of the words composing the list was presented to the participants (see Appendix 6). According to the approach instruction condition, participants were either told that they would have to approach the Luupite and avoid the Niffite or the opposite. Participants were told that these instructions were relevant for the two days of the experiment. As in Experiments 5–7, these instructions were described as critical for the experiment. Because instructions would not be presented again, participants had to correctly memorize them.

While participants in the approach-avoidance instructions directly completed an IAT after receiving the instructions, participants in the approach-avoidance training did not (see Figure 9.2, p. 183). The procedure that we used for this experiment was similar to the one we used when we tested the online version of the approach-avoidance training (Experiment 1b). The approach-avoidance training was implemented within the VAAST. Each trial started with a fixation cross followed by a target word at the center of the screen. Participants could either perform an approach movement or an avoidance movement by pressing twice the directional arrow keys of their keyboard—each press triggered an animation consistent with the movement. Errors were indicated by a red cross, and participants were invited to start the trial over again. As in Experiment 1b, participants started by a 20-trial practice block and then moved to a 200-trial main block; one half of the trials was related to approach, and the other one was related to avoidance. The approach-avoidance training phase lasted

approximately 10 minutes. After the approach-avoidance training, participants from the approach-avoidance training condition completed an IAT.

Both participants in the approach-avoidance training condition and the approach-avoidance instructions completed an identification IAT identical to the one included in Experiment 1b. In this task, participants categorized words from the Luupite and Niffite lists, and words related to the self and others using the “E” and “I” keyboard keys from (for a list of the stimuli, see Appendix 6). Participants performed this categorization task in several configurations, the critical ones being when they had to categorize the two types of categories at once (see Figure 9.3). In one of the critical blocks, participants categorized Luupite and self-related words with a single response key (e.g., the E key) and Niffite and others-related words with another single response key (e.g., the I key). In the other critical block, participants categorized Niffite and self-related words with a single response key and Luupite and others-related words with another single response key. As in Experiment 1b, we recorded response latencies from the 20-trial critical practice blocks and the 72-trial critical test blocks. The location of the key associated with the self, as well as the key associated with the Luupite category in the first part of the IAT was counterbalanced across participants and stayed the same during Stage 2.

The IAT was followed by an attention check. Participants were asked to “report their favorite color [based on the text above]”. The said text was asking them to write “croissant” when asked for their favorite color. Participants answered this question, and then reminded that a second stage of the experiment would happen the day after. Participants were thanked for their participation.

Figure 9.3*IAT Block Configuration in Experiment 6 (Stage 1 & Stage 2)*

	Category label	Target word	Category label
Configuration 1	self		other
- Practice block	E	myself	I
	E	them	I
Configuration 2	Luupite		Niffite
- Practice block	E	Maasolup	I
	E	Cellanif	I
Configuration 3	self		other
- Practice block	Luupite		Niffite
- Test block	E	myself	I
	E	them	I
	E	Maasolup	I
	E	Cellanif	I
Configuration 4	Niffite		Luupite
- Practice block	E	Maasolup	I
	E	Cellanif	I
Configuration 5	self		other
- Practice block	Niffite		Luupite
- Test block	E	myself	I
	E	them	I
	E	Maasolup	I
	E	Cellanif	I

Note. The key associated with the self and the key associated with Luupite in the first part of the IAT was counterbalanced across participants.

Stage 2. The second stage of the experiment was made available to the participants the day after, at 7:00 GMT+1. This ensured us that at least seven hours elapsed between Stage 1 and Stage 2. We made Stage 2 available to our participants through the Prolific

platform (Palan & Schitter, 2018). Participants were welcomed back and informed that this stage of the experiment would last 10 minutes. As in Stage 1, the experiment was set to Fullscreen on participants' computers. Participants were reminded of Luupite and Niffite stimuli, but not the movements associated with the stimuli.

The first task participants had to complete was a categorization task. It was the same identification IAT as the one used in Stage 1. The key associated with the self category and the key associated with the Luupite category in the first part of the experiment was the same as the ones in Stage 1.

Right after, as in Stage 1, we included an attention check. Participants were asked to write "croissant" when asked for their favorite color. They were then debriefed and thanked for their participation in the experiment.

Results and Discussion

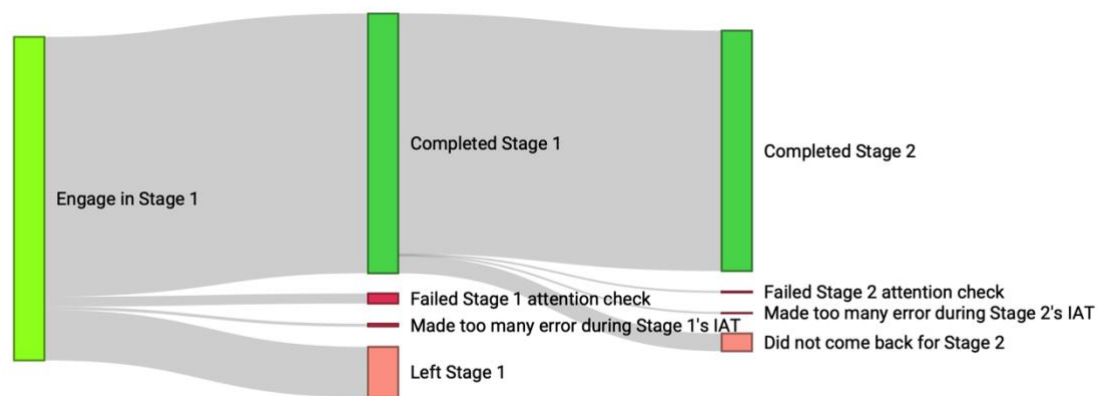
In this experiment, we wanted to test whether the effects of approach-avoidance training and approach-avoidance instruction were the same over time. To do so, we investigated whether the difference of evaluation between (to be) approached and (to be) avoided stimuli was moderated by the type of training and the stage of the experiment. We expected the decrease over time in the difference of evaluation between (to be) approached and (to be) avoided stimuli to be of lower magnitude when participants completed an approach-avoidance training compared to when they received approach-avoidance instructions.

We did not include the data from every participant in our confirmatory analysis (see Figure 9.4). The exclusions that we made were preregistered. A total of 463 participants completed the first stage of the experiment. We excluded 17 participants who failed the attention check (3.67%) and 5 participants who had an error rate over 30% during the IAT (1.08%). Out of the 441 participants who were eligible for Stage 2, 411 took part in the

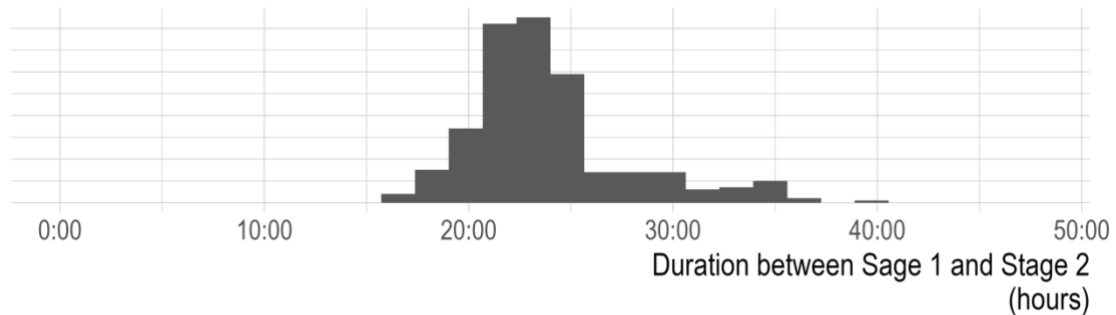
second stage of the experiment—30 participants did not come back (6.80%). Among the participants who completed Stage 2, we had to exclude 2 participants who failed the attention check (0.49%) and 1 participant who had an error rate over 30% during the IAT (0.24%). In total, data from 408 participants were included in the analysis. On average, participants took part in the second stage of the experiment 23 hours and 55 minutes ($SD = 4.16$ h) after completing the first stage (see Figure 9.4).

Figure 9.4

Flowchart of Participants Engaging with Stage 1 in Experiment 6



Note. Participants included in the confirmatory analysis are participants who successfully completed the Stage 2 experiment ($n = 408$).

Figure 9.5*Duration Between Stage 1 and 2 in Experiment 6*

As preregistered, we conducted a mixed-effect ANOVA to test our hypotheses.³¹ As in Experiment 1b, which also was an online experiment, we used the D2 IAT score as the outcome (Greenwald et al., 2003). This score was computed so that a positive score indicates that a participant is faster in the block which requires the self- and the Luupite-related stimuli with the same response key. We used the approach instructions (approach Luupites vs. approach Niffites) and the type of approach-avoidance procedure (approach-avoidance training vs. approach-avoidance instructions) as between-subject predictors and the time of the measure (Stage 1 vs. Stage 2) as a within-subject predictor. Theoretically relevant effects are the ones qualifying the effect of approach instructions. For readability reasons, we will only report these ones in the text, but a full summary of the ANOVA statistics can be found in Table 9.1.

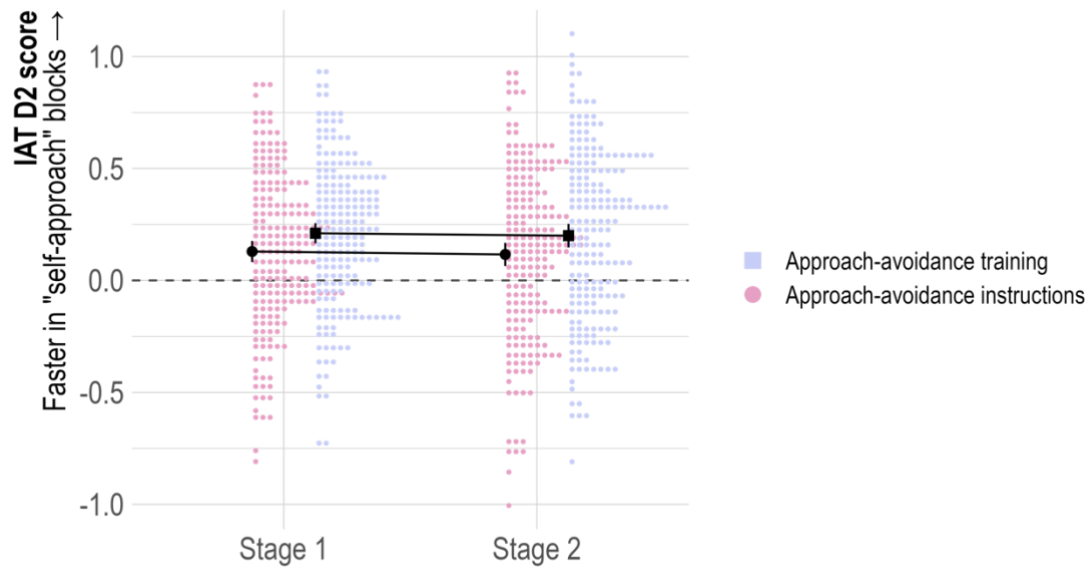
³¹ Note that these results could have been analyzed with mixed models (Wolsiefer et al., 2017). However, at the time we preregistered this experiment, we preferred to use the same type of analysis for all of our online experiments. In other words, we wanted to keep an analytical strategy similar to the one used in Experiment 1b.

First, our analysis revealed that, on average, participants were faster to perform the IAT in blocks where the self and approached stimuli shared a response key, $F(1, 404) = 107.43, p < .001, \eta_p^2 = .210$. Importantly, across the two stages of the experiment, this effect was stronger when participants performed an approach-avoidance training compared to when they received approach-avoidance instructions, $F(1, 404) = 6.58, p = .010, \eta_p^2 = .016$. This difference of effects between the two procedures did not vary significantly between the two stages of the experiment, $F(1, 404) = 0.01, p = .942, \eta_p^2 < .001$.³² Regarding the effect of time, on average, the difference of evaluation between the two kinds of stimuli did not differ significantly between the two stages, $F(1, 404) = 0.45, p = .446, \eta_p^2 = .001$ (see Figure 9.6).

³² Obviously, this result can also be interpreted in terms of stability of evaluation not significantly differing across the two approach-avoidance procedures.

Figure 9.6

Effects of Approach-Avoidance Procedures Over Time (Experiment 6)



Note. Each dot represents a participant. The error bars represent 95% CI. The error bars related to the approach-avoidance procedures are aligned with the participants from their condition.

Table 9.1*Summary Statistics of Experiment 6's Model*

	Estimate	95% CI	F	p-value	η_p^2
Approach instructions	0.33	0.26; 0.39	107.43	< .001	.210
AA procedure	-0.03	-0.10; 0.39	1.17	.280	.003
Stage	-0.01	-0.04; 0.02	0.265	.607	< .001
Approach instructions × AA procedure	0.16	0.04; 0.29	6.58	.010	.016
Approach instructions × Stage	0.02	-0.09; 0.04	0.45	.446	.001
AA procedure × Stage	0.00	-0.06; 0.06	< 0.01	.984	< .001
Approach instructions × AA procedure × Stage	0.00	-0.12; 0.13	0.01	.942	< .001

Note. Estimates are based on a model where the approach-avoidance procedures and the approach-instructions are contrast-coded (approach-avoidance instructions = -0.5; approach-avoidance training = 0.5; approach Niffites = - 0.5; approach Luupites = 0.5).

For the first time since we started investigating the difference between approach-avoidance trainings and approach-avoidance instructions, our analysis revealed a significant difference between these two types of procedure. Our reasoning had led us to believe that a difference between the two procedures would be especially likely to appear during the second stage of this experiment. Yet, we observed a superiority of approach-avoidance training compared to mere instructions during the first stage of the study, $t(404) = 2.39$, $p = .017$, $d_s = 0.24$, 95% CI [0.04, 0.43]. These results could be considered unexpected given our previous results (Experiments 5–7). As a consequence, we investigated two hypotheses regarding the statistical procedures that we used which could account for the differences in significance of the difference between the two types of procedures between Experiments 3–5

and Experiment 6. Beforehand, however, there are few differences that we want to highlight between Experiments 3–5 and Experiment 8. While we will not be able to test whether the difference of significance between the two approach-avoidance procedures across our experiments is due to these factors—and, in fact, we will not be able to test whether the difference between these experiments is significant—we think that it is important to keep in mind.

First of all, an important difference between these experiments is the evaluative measures used. In Experiments 3–5, we adopted an AMP (Payne et al., 2005); In Experiment 8, we adopted an IAT (Greenwald et al., 1998). This difference between the procedure used highlights one limitation with the models of evaluations we focus on (i.e., the SEM, the APE model, or single-process propositional models of evaluation), that is, the low formalization of the processes that might have some importance during evaluative response expression. Notably, these models do not distinguish the processes involved in the different indirect evaluative response. Another difference would be the type of participants that we used. In Experiments 3–5, we recruited undergrad psychology students; In Experiment 6, we recruited participants from the general population. This factor could be considered of importance because models such as Van Dessel's (2018a) single-process propositional model suggest that evaluation acquisition should be sensitive to individual differences and undergrad students are thought to differ from the characteristics of the general population (Sears, 1986). Still, at this point, any consideration regarding the differences between Experiments 3–5 and Experiment 6 should be considered speculative. We will now turn into the statistical reasons why we could have observed nonsignificant results in Experiments 3–5 and significant results in Experiment 6.

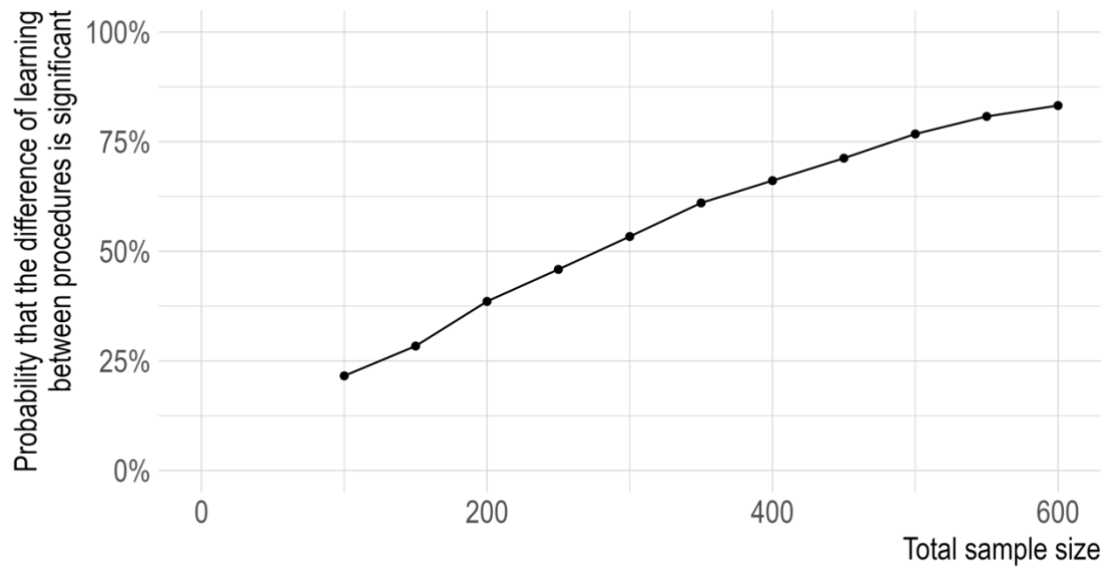
A first explanation could be that the significant difference between the approach-avoidance procedures that we observed during the first stage of Experiment 6 is a false

positive. While this hypothesis is not the most inspiring, theoretically speaking, it still is a possibility. In the present dissertation, in terms of data analysis, we adopted a frequentist approach with a nominal alpha rate of 5%. In Experiment 6, we preregistered our analysis with a level of specificity which, we hope, prevent flexibility in data analysis. We did not deviate from this preregistered plan and the criteria for stopping data collection was set a priori. This strategy makes it unlikely that we increased the risk of false positive to a factor higher than 5% (Simmons et al., 2011).

A second hypothesis is the inverse of the first one. Rather than a false positive in Experiment 6, we could have observed false negatives in Experiments 3–5. In other words, the difference of effect between the two approach-avoidance procedures might be too small for us to have been able to reliably observe it. This assumption is consistent with the fact that we observed a significant difference between the approach-avoidance procedures in Experiment 6, an experiment with a sample size substantially larger than the one used in Experiments 3–5. This assumption is also consistent with a bootstrapping analysis that we conducted showing the results of Experiment 6 are very sensitive to the sample size used (see Figure 9.7).

Figure 9.7

Difference Between the Approach-Avoidance Procedures' Sensitivity to Sample Size



Note. We estimated the proportion of experiments which would find a significant difference between approach-avoidance training and approach-avoidance instructions procedures in terms of evaluative learning using bootstrap methods. For each of the sample size that we computed, we performed 5000 resamples.

Overall, the different post-hoc hypotheses that we generated to account for our inconsistent results regarding the difference between approach-avoidance training and approach-avoidance instructions in terms of evaluative learning highlights the usefulness of a highly powered replication. This is what we decided to do in Experiment 7.

Experiment 7

In Experiment 6, for the first time since we began this line of research, we observed a significant difference between approach-avoidance training and approach-avoidance instructions. None of the experiments that we conducted, beside this one, indicated a

significant difference between the two procedures in terms of evaluative learning. The goal of the present experiment is to investigate the robustness of the difference observed in Experiment 6. To do so, we will replicate the first stage of this previous experiment. Precision in effect size estimation will be critical for the present experiment. We report all measures, manipulations, and exclusions. Material, data, the analysis script, the analysis script, and the preregistration can be found online at <https://osf.io/t4w3e>.³³

Method

Participants, Design, and Power Analysis

In this experiment, we pre-registered a sample size of 1000 participants. In total, 996 participants ended up completing the experiment in exchange for £1.80 ($M_{age} = 36.59$, $SD_{age} = 13.67$; 536 females, 429 males, 31 unreported gender). Participants were recruited using the Prolific platform (Palan & Schitter, 2018). To take part in this experiment, they must be native English speakers, had to be British, had an approval rate over 90% on the Prolific platform, and it had to be the first time that they participated in one of our studies.

Participants took part in an experiment with a one between-subject experimental design (approach-avoidance training vs. approach-avoidance instructions). The stimuli participants had to approach (Luupite vs. Niffite), the location of the self category during the IAT (left vs. right), and the location of the approached stimulus during the first block of the IAT (left vs. right) were counterbalanced across participants.

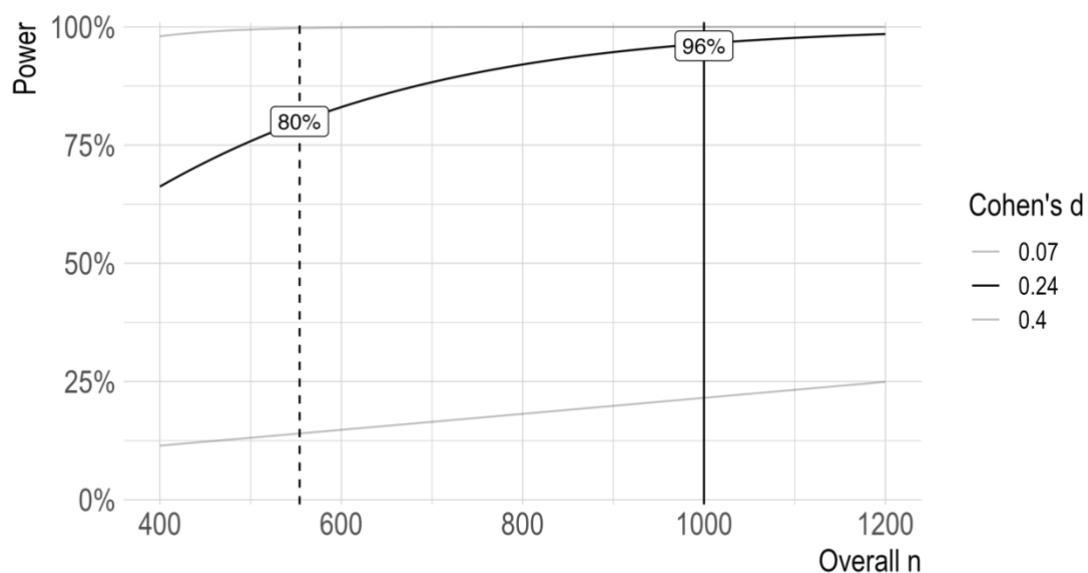
Because precision was critical for this experiment, we conducted a power analysis. In the first stage of Experiment 6, the evaluative response acquired was more pronounced when

³³ This repository might be subject to an embargo. If it is unavailable, please e-mail cedric.batailler@gmail.com.

participants approached and avoided the target compared to when they merely received approach-avoidance instructions, $d_s = 0.24$ CI95% [0.04, 0.43]. Achieving .80 power for a $d_s = 0.24$ effect requires recruiting 544 participants. By recruiting 1000 participants, one would obtain .96 statistical power for a $d_s = 0.24$ effect size (see Figure 9.8). We decided to adopt this sample size. This decision was based on our objectives with this experiment, resource constraints, and feasibility.

Figure 9.8

Sensitivity Analysis for Experiment 7



Note. Statistical power as a function of sample size for several effect sizes. $d_s = 0.34$ corresponds to the effect size of the difference between the two approach-avoidance procedures in terms of evaluation acquisition in Experiment 6 (Stage 1). $d_s = 0.07$ and $d_s = 0.40$ corresponds to the lower and higher bounds of the 90% CI of this effect.

Procedure and Material

As in Experiment 6, participants were recruited on the Prolific platform (Palan & Schitter, 2018). They were informed that they would participate in a study which could last up to 20 minutes. The study could include attention checks. Upon connection to the experiment, participants were randomly assigned to the approach-avoidance training or the approach-avoidance instructions condition. As in our other online experiments (Experiments 1b, 3, 4, & 8), the experiment was coded using jsPsych (de Leeuw, 2015) and was set automatically on Fullscreen.

Participants first received instructions related to the Luupite or the Niffite categories. Participants would have to approach the Luupites and avoid the Niffites one (or the opposite) and that they would receive instructions on how to proceed later. Participants in the approach-avoidance training condition then moved on to the training. This approach-avoidance training was identical to the one included in Experiment 6. That is, in a VAAST-based procedure, participants started by a 20-trial practice block and then moved to a 200-trial main block. The approach-avoidance training lasted approximately 10 minutes.

Both participants in the approach-avoidance training and approach-avoidance instruction then completed an IAT. As in Experiment 6, participants categorized words related to the self, to others, to the Luupites, and to the Niffites. The key associated with the self category and the key associated with the Luupite category in the first part of the experiment was counterbalanced across participants.

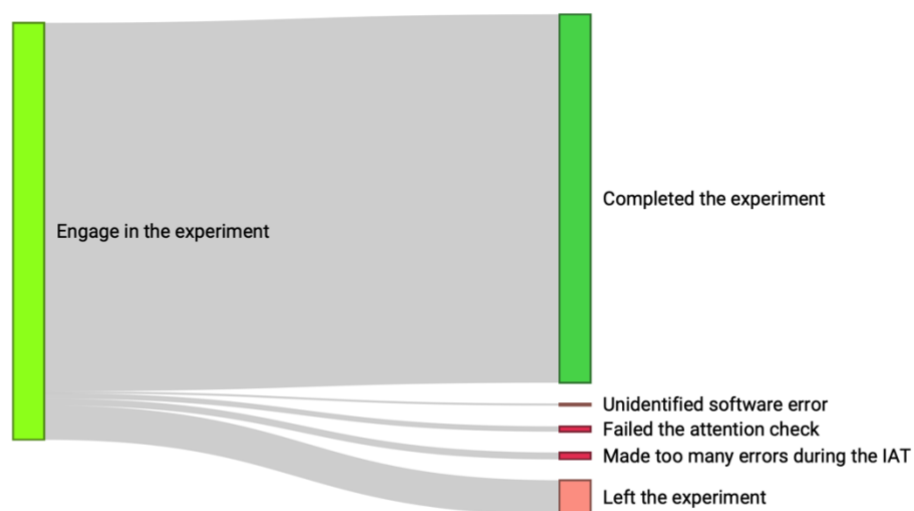
We included an attention check after the participants performed the IAT. When asked for their favorite color, participants had to write the word “baguette” instead. At the end of the experiment, participants were thanked and debriefed.

Results and Discussion

Out of the 1091 participants engaging with the experiment, 90 left before completion. We lost data from 5 participants because of unidentified software errors.³⁴ Out of the 996 remaining participants, as preregistered, we excluded 14 participants because of the failed attention check (1.4% of the participants) and 18 participants because they made too many errors during the IAT (1.8% of the remaining participants, i.e., more than 30% of the trials were incorrect). In total, data from 964 participants were included in the analysis (see Figure 9.9).

Figure 9.9

Trajectory for Every Participant in Experiment 7

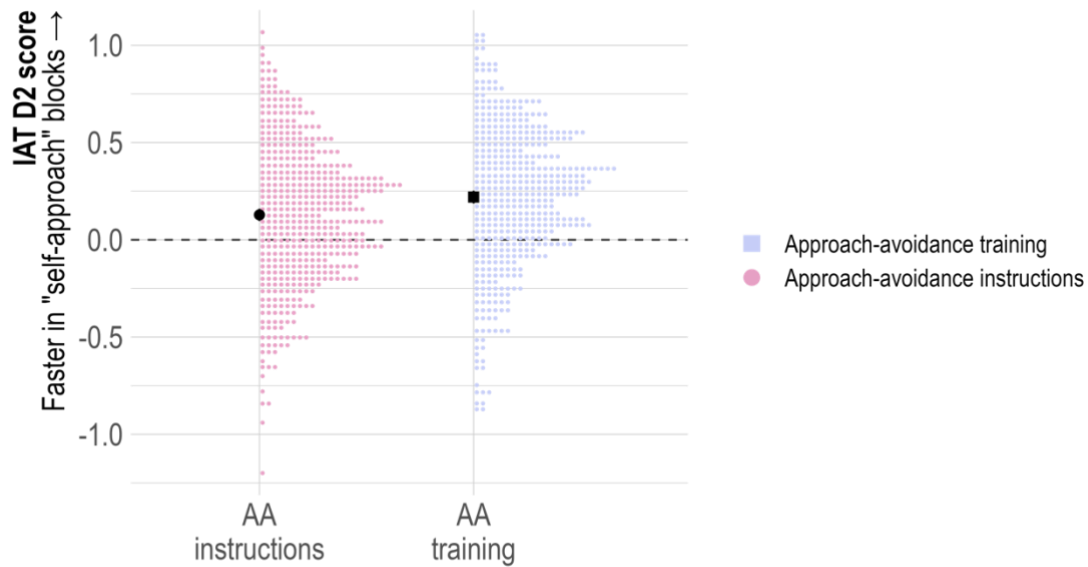


³⁴ The error resulted in one participant completing the study twice and four participants for which we could not recover the data related to the attention check.

In this experiment, as preregistered, we adopted the D2 scoring algorithm for the IAT: We excluded trials where the response time was below 400 ms or above 10000 ms and we included error latencies in the analysis (i.e., we did not exclude incorrect trials). We standardized the difference of time latencies between the critical blocks by dividing this difference by the standard deviation of the of all the latencies in the two test blocks. Unlike Experiment 6, we computed the difference, so that a positive score indicated that participants were faster un the block where the (to be) approached category and the self shared a response key (rather than so a positive score indicates that participants are faster in the self-Luupite block). This way to compute the D2 score allow us to investigate whether participants acquired an evaluative response by assessing whether the scores are significantly positive. We used this score as a dependent variable.

To determine the effectiveness of the approach and avoidance training and approach and avoidance instructions in forming evaluative responses, we analyzed this dependent variable by adopting a model comparison approach with the type of training as a predictor (actual training vs. instructions).³⁵ This analysis first revealed that, for an average effect of the training, participants were faster in the block where the (to be) approached stimulus and the self shared a response key compared to when they did not, $t(962) = 15.00, p < .001, d_z = 0.48, CI95\% [0.41; 0.55]$. Importantly, this analysis also revealed that this difference was larger when participants actually performed the training compared to when they merely received the approach-avoidance instructions, $t(962) = 3.94, p < .001, d_s = 0.25, CI95\% [0.13; 0.38]$.

³⁵ This analysis is the same than a *t*-test for independent samples, but allows us to investigate whether the

Figure 9.10*Evaluative Learning According to Approach-Avoidance Procedures*

Note. Each dot represents a participant. Error bars (albeit small) represent the 95% CI.

AA = Approach-avoidance.

We wanted to investigate whether the difference in terms of evaluative learning between the two approach-avoidance procedures that we observed at the first stage of Experiment 6 was significant again in a highly powered experiment. Experiment 7 revealed that it was. Indeed, in this experiment, we observed that both approach-avoidance training and approach-avoidance instructions resulted in the formation of an indirect evaluative response on an IAT, and that the evaluative responses acquired as a result of approach-avoidance training were more pronounced. The superiority of approach-avoidance training is a small size effect, $d_s = 0.25$, 95% CI [0.13; 0.38]. This suggests that our result in Experiment 6 was not a false positive.

General Discussion

In this chapter, we investigated the evaluative learning resulting from approach-avoidance training and from approach-avoidance instructions over time. In Experiment 6, our participants first performed an approach-avoidance training or received approach-avoidance instructions. The identification IATs that participants completed right after the procedures and the day after allowed us to investigate how the acquired evaluative responses evolve over time. We expected that evaluative responses acquired from an approach-avoidance training would last longer than the one acquired through approach-avoidance instructions. This experiment did not allow to find a difference between the two procedures regarding the sustainability of the indirect evaluative responses the two procedures resulted in. We found no evidence that actual approach-avoidance training (compared to mere approach-avoidance instructions) produced more durable indirect evaluative response; however, we found that the evaluative responses acquired as the result of the training were more pronounced than the one acquired as a result of approach-avoidance instructions. This last result was replicated in Experiment 7.

When conducting Experiment 6, we made the hypothesis that operational differences between approach-avoidance training and approach-avoidance instructions could impact the likelihood that the relevant information resulting from an approach-avoidance procedure was encoded from working memory to long-term memory. If so, we should have observed a difference in the decrease of the strength of indirect evaluative responses between the two days of this experiment. Our results did not provide support for this hypothesis. Indeed, a first observation we can make out of Experiment 7 is that the indirect evaluative responses acquired through approach-avoidance lasted at least until the next day. Moreover, we did not find evidence that the acquired evaluative responses decreased in strength over time. Because it is highly unlikely that the propositions related to the (to be) approached and avoided stimuli

stayed in working memory in between two days, these results suggest that the information impacting evaluation has been encoded in a longer-term memory system. These effects are consistent with other experimental work showing that a formed indirect evaluative response last in time. For example, Förderer and Unkelbach (2013) showed that the indirect evaluative responses acquired for neutral stimuli as a result of an Evaluative Conditioning procedure could last up to one week. These data suggest that the acquisition of an indirect evaluative response is not limited to the close instants following the acquisition procedure.

The literature on the lasting effect of procedures targeting indirect evaluative responses, however, highlights an important distinction. Evaluative response acquisition refers to two kinds of situations: evaluative response formation and evaluative response change. While the former describes a situation in which new experiences impact the evaluative responses toward a stimulus which did not initially evoke an evaluative response, the latter describes a situation in which new experiences impact the evaluative response toward a stimulus which did evoke an evaluative response. This distinction often overlaps with the kinds of stimuli targeted by the training: novel (e.g., Luupite and Niffite) or not (e.g., math or arts). In a study investigating the effects of nine different interventions targeting indirect evaluative responses related to Black and White people, Lai et al. (2016) could not provide evidence that the changes in indirect evaluative responses lasted over time. As we will see, these observations related to the long-term acquisition of evaluative responses are consistent with the assumptions regarding the reactivation of the relevant proposition in time (De Houwer, 2018).

The single-process propositional model of evaluation (De Houwer et al., 2013) suggests that the similarity between a subjective experience and encoded episodic memory traces will impact in the reactivation of memory traces. The more similar the traces are to the subjective reality, the higher the likelihood that the traces will be reactivated. Once the traces

reactivated, they will influence evaluation processes. With this framework in mind, it is possible to make different predictions regarding the effects of approach-avoidance procedures and their longer-term impact on indirect evaluative response. In the case of evaluative response formation, it is assumed that the stimuli targeted by the procedure were not present in memory traces before. A consequence is that the only kind of traces likely to be reactivated after an approach-avoidance procedure is the one containing information related to approach and avoidance. The same is not true when it comes to indirect evaluative response change, where the traces likely to be reactivated are not only the ones encoded during the training. Besides these considerations on the long-term effects of approach-avoidance training, Experiment 6 was instructive regarding the difference of strength between the evaluative responses acquired through training and through instructions.

Compared to our previous experiments (Experiments 3–5), Experiment 8 revealed a difference between the evaluative responses acquired as a result of an approach-avoidance training or approach-avoidance instructions. Specifically, the difference of evaluation between (to be) approached stimuli and (to be) avoided one was larger when participants actually performed the approach and avoidance movement compared to when they merely received approach-avoidance instructions (as measured with an identification IAT; Nosek et al., 2002). This result was confirmed by a replication of the first stage of Experiment 8 in Experiment 7. In Experiments 6 and 7, the difference of effects between the two procedures was a small size effect ($d_s = 0.25$; Cohen, 1988). At first, these results might appear inconsistent with our previous experiments, but the sample sizes adopted Experiments in 3–5 indicate that our previous experiments were actually underpowered for an effect size that small. In other words, if the difference between the two procedures in terms of effect size was the same across every experiment where we compared the two procedures (i.e., Experiments 5–9), we would not have been able to reliably observe it in Experiments 3–5.

From a theoretical perspective, the observation of this difference aligns with the idea that repeated experience contributes to indirect evaluation. This idea initially appeared in the early days of implicit social cognition (Fazio et al., 1986; Greenwald & Banaji, 1995) and is now retrieved in modern theories of evaluation (e.g., the SEM, the APE model, or the single-process propositional model). The different models that we reviewed in the introduction section of this manuscript make specific predictions regarding the role of repeated experience, regardless of whether the model use associative processes to account for indirect evaluative response acquisition. With the present experiments, we provide support for an added value of repeated experience in the learning of indirect evaluative responses. This added value, however, appears to be of a small effect size.

Conclusion

In this chapter, in a highly powered experiment (Experiment 6), we investigated the effects of approach-avoidance procedures after a delay. The key findings we observed were that the approach-avoidance training outperformed approach-avoidance instructions when it comes to the acquisition of an indirect evaluative response and that the effects of both procedures lasted in time. We did not find evidence for a decrease in strength of the acquired evaluative responses over time, that is neither for the approach-avoidance training nor the approach-avoidance instructions conditions. The difference between approach-avoidance training and approach-avoidance instructions was replicated in Experiment 7.

Chapter 10

Lay Beliefs About Approach-Avoidance Procedures

In this dissertation, our objective was to investigate the role of repeated experience when it comes to the acquisition of indirect evaluative response. From an operational perspective, this will lead us to compare the effects approach-avoidance training and approach-avoidance instructions. In the previous chapter, we showed that the evaluative responses acquired as a result of actual approach-avoidance training could be more pronounced than those acquired as a result of mere approach-avoidance instructions (Experiments 6 & 7). Yet, it seems that the difference between the two procedures could be a small effect size. Given the range of application of approach-avoidance procedures, this suggests that approach-avoidance training procedures (over mere instructions) are relevant. In the present chapter, we investigated a factor which could have its importance: the lay beliefs of participants regarding the effectiveness of approach-avoidance procedures.

The Effects of Approach-Avoidance Training and Approach-Avoidance Instructions

As we saw when reviewing the implicit social cognition literature dedicated to this phenomenon, most models assume that repeated experience will result in something special in the human mind. For the SEM, repeated experience is a *sine qua non* condition for the emergence of new indirect evaluative responses (McConnell & Rydell, 2014; Rydell & McConnell, 2006). For the APE model, new indirect evaluative responses can emerge without someone repeatedly experiencing something, as propositional processes can interact with associative ones (Gawronski & Bodenhausen, 2006, 2007, 2014a). The acquisition of a proposition can ricochet in the acquisition of an indirect evaluative response through associative processes. Still, when a two diverging evaluative responses exist, one resulting from direct associative processes and the other resulting from an indirect route originating by

propositional processes, the APE model assume that the direct one will override the other (Hu et al., 2017). From the perspective of single process propositional models (De Houwer, 2009, 2018; Van Dessel et al., 2018a), because the acquisition of a proposition (which ultimately influences indirect evaluative responses) is an effortful process, repetition makes it more likely for someone to acquire a proposition, and, therefore, an indirect evaluative response. In the end, every model assumes a specific role of repeated experience.

As discussed in the introduction, some predictions of the models appeared during the early days of implicit social cognition research. At the time, researchers developed theories and tools to capture something so strongly rooted in participants mind that it would be observable even if participants would not want to disclose it or even if participants were unaware of it (Payne & Gawronski, 2010). In the present chapter, we investigate whether lay theorists—unaware of the challenges of implicit social cognition—make the same assumptions as the implicit social cognition researchers, have regarding the role of repeated experience when it comes to the acquisition of indirect evaluative response.

On the Importance to Investigate People's Lay Belief

At first glance, investigating people's lay belief about how repeated experience could impact them could seem like a curious choice, especially given that evidence showing the limits of human introspection exist in the literature (e.g., Nisbett & Wilson, 1977). One of the reasons why the question of the lay beliefs of participants is important lies in the applications of approach-avoidance training. One of them is clinical. When tackling addictions, most theoretical models of evaluation suggest that past experience with a stimulus (e.g., cigarette) will impact behavior in subsequent encounter with the stimulus (De Houwer, 2018; Gawronski & Bodenhausen, 2007; Rydell & McConnell, 2006). For example, because of their experience with cigarettes, the mere perception of a cigarette for a smoker could trigger an approach bias (C. E. Wiers et al., 2013). Because such automatically emitted behavior

could represent an obstacle when people try to quit smoking, approach and avoidance training have been used as add-on to therapies. The idea here is to override problematic automatic responses by making individual acquire new ones (for recent reviews and meta-analysis (Boffo et al., 2019; Kakoschke et al., 2017)).

Approach-avoidance training quickly showed their efficacy (R. W. Wiers et al., 2011). Given that empirical data (Van Dessel et al., 2015) and some theoretical models of evaluation (Gawronski & Bodenhausen, 2007; Van Dessel et al., 2018a) suggest that approach-avoidance instructions could also be effective, investigating their effects in this context could be interesting, especially because they are easy to implement. A factor, however, could limit approach-avoidance instruction-based intervention. Just like McConnel and Rydell's SEM (2014), people could think that these procedures are not effective. Outcome expectancies refers to the prospective belief a patient can have regarding the effectiveness of a therapy or an intervention (Scheier & Carver, 1985). This factor has been linked to psychotherapy outcomes in several ways, usually showing that psychotherapy has better effect among people who are confident on their effects (for a review, see Greenberg et al., 2006). Because of the importance of this factor in an intervention's effectiveness, we consider the question of the lay beliefs people hold regarding approach-avoidance intervention relevant.

People's Lay Theories Regarding Approach-Avoidance Procedures

In approach-avoidance training, participants receive approach-avoidance instructions regarding a type of stimulus. Most of the time, the instructions are related to the stimuli of

which researchers are trying to change the indirect evaluation.³⁶ That is, when researchers are interested in testing whether an approach-avoidance training can change evaluation of cigarettes, they will ask their participants to avoid pictures related to cigarettes and approach pictures not related to cigarettes (but see Boffo et al., 2017). For this reason, one can consider that *approach-avoidance training* includes *approach-avoidance instructions*. In other words, if approach-avoidance instructions are a and the experience of approach-avoidance experiences is b , approach-avoidance training is $a + b$.

Because approach-avoidance training contains approach-avoidance instructions, it is likely that naive participants make the same prediction as implicit social cognition theories. Approach-avoidance training should outperform approach-avoidance instructions when it comes to the acquisition of an evaluative response. We tested whether it was really the case in Experiments 8 and 9. In these two preregistered experiments, participants had to make predictions regarding the outcomes of two interventions whose purpose was to form indirect evaluative responses: approach-avoidance trainings and approach-avoidance instructions. Participants from Experiments 8 and 9 were asked how they think these interventions would affect them. The two experiments were very similar, except for an approach-avoidance training demo included in the Experiment 9.

Experiment 8

The objective of Experiment 8 was to assess whether people thought that approach-avoidance training was more effective than approach-avoidance instructions at forming indirect evaluative responses. This experiment was preregistered. We report all measures,

³⁶ Note that researchers are seldom interested in changing indirect evaluations (e.g., an attitude toward cigarette IAT) rather than problematic behaviors.

manipulations, and exclusions. Material, data, the analysis script, the analysis script, and the preregistration can be found online at <https://osf.io/he2qu/>.³⁷

Method

Participant, Design, and Sensitivity Analysis

One hundred and fifty participants took part in this experiment in exchange for £0.70 ($M_{\text{age}} = 37.5$, $SD_{\text{age}} = 11.9$; 78 females, 72 males). Participants were recruited on the Prolific platform (Palan & Schitter, 2018). They had to be native to the UK or the US and they had to have an approval rate over .90 on the Prolific platform to engage with our experiment. We used this inclusion criteria as proxies for their proficiency with English and seriousness when engaging with online experiments. These two factors were important as this experiment required participants to read a lot of text.

We decided to adopt a design with one within-participant variable for this experiment: the type of training to evaluate. Participants taking part in the experiment therefore had to evaluate the efficacy of both approach-avoidance training and approach-avoidance instructions. The order of evaluations was counterbalanced with random assignment.

We preregistered a sample size of 150 participants, which is sufficient to detect small effect sizes (Cohen, 1988). The sensitivity analysis that we conducted revealed that we had .79 statistical power to detect a small effect size (i.e., $d_z = 0.20$), over .99 statistical power to detect a medium effect size (i.e., $d_z = 0.50$), and over .99 statistical power to detect a large effect (i.e., $d_z = 0.80$). Given that we have around .80 statistical power to detect what is

³⁷ This project might be subject to an embargo. If it is unavailable, please e-mail cedric.batailler@gmail.com to gain access to this project.

traditionally considered a small effect size in psychology (Cohen, 1988), we considered our experiment adequately powered.

Procedure and Material

Participants were recruited from their web browser thanks to the Prolific platform (Palan & Schitter, 2018). From there, participants were redirected to a Qualtrics page where we informed them that they had to imagine the outcome of two interventions. The two types of intervention were aiming at “creating attitudes for two fictitious groups”.

One challenge of this study was to provide participants with enough details for the interventions to appear meaningful. We suspected that a description of the studies as theoretically agnostic as the ones that we use in the present dissertation (i.e., procedures aiming at changing *indirect evaluative responses*) would make it harder for the participants to engage with the material. Therefore, we decided to describe attitudes as “an inclination to evaluate things positively or negatively” and told participants that the researchers were capturing it with an IAT, a tool measuring “the spontaneous reaction [a stimulus] evokes” (see Appendix 7 for the verbatim of Experiment 3).

In this experiment, the two fictitious interventions were aiming at forming attitudes toward Luupite and Niffite social groups (see Gregg et al., 2006). Once this goal was made explicit, descriptions of the two types of intervention were provided. According to the counterbalancing order, participants first read the description of the approach-avoidance training intervention or the approach-avoidance instructions. The approach-avoidance training intervention was described as follows: “In the “training” condition, [as a participant] you would receive instructions indicating that in a task that will follow, you will have to approach members of the Niffite group and avoid members of the Luupite group. The researchers would then tell you how to proceed: Group members’ names will appear in a virtual environment like in a video game representing a street. In this environment, you will

have to approach and avoid members as previously instructed, by using the arrow keys on your keyboard. You would then actually complete the training which consists of 200 trials; it would last approximately 10 minutes”. The approach-avoidance instructions intervention was described with the following terms: “In the “instructions” condition, [as a participant] you would receive instructions indicating that in a task that will follow, you will have to approach members of the Niffite group and avoid members of the Luupite group. The researchers would explain to you that you will receive complementary information on how to approach and how to avoid the group members later on, but first you would have to go through another unrelated task. Thus, you would not actually perform the approach and avoidance movements toward the group members”. Both descriptions ended indicating that participants of these fictitious experiment had to perform an IAT after the intervention. This task was described as “[measuring] your implicit attitudes toward Luupite and Niffite, that is, the spontaneous affective reactions these groups evoke”. A verbatim of the Experiment 3a is available in Appendix 7.

After reading the descriptions, participants made predictions regarding the outcome of the two interventions. The order of the predictions was the same as the descriptions. To make their prediction regarding the outcomes of the interventions, they had to answer the following question with a visual analog scale: “If you were a participant in the [training/instructions] condition, how much do you think your Implicit Association Test score would be influenced by [this training/these instructions]?”. The minimum and maximum of this scale respectively mapped onto the opposite of the intended effect and the intended effect. The scale ranged from a minimum value indicating that the intervention would create a negative attitude for the (to be) approached stimuli and a positive attitude for the (to be) avoided stimuli to a maximum value indicating that the intervention would create a positive attitude for the (to be)

approached stimuli and a negative attitude for the (to be) avoided stimuli. This scale also had an explicit mid-point indicating that the intervention would create no attitudes (see Figure 10.1).

After completing the experiment, participants were thanked for their participation, debriefed, and paid. The whole study lasted approximately seven minutes.

Figure 10.1

The Measure of Expected Effectiveness of the Interventions, as Measured in Experiment 9

If you were a participant in the **training condition**, how much do you think your Implicit Association Test score would be influenced by this training?



If you were a participant in the **instruction condition**, how much do you think your Implicit Association Test score would be influenced by receiving these instructions?



Note. Order of the questions were counterbalanced to match the order of the intervention descriptions.

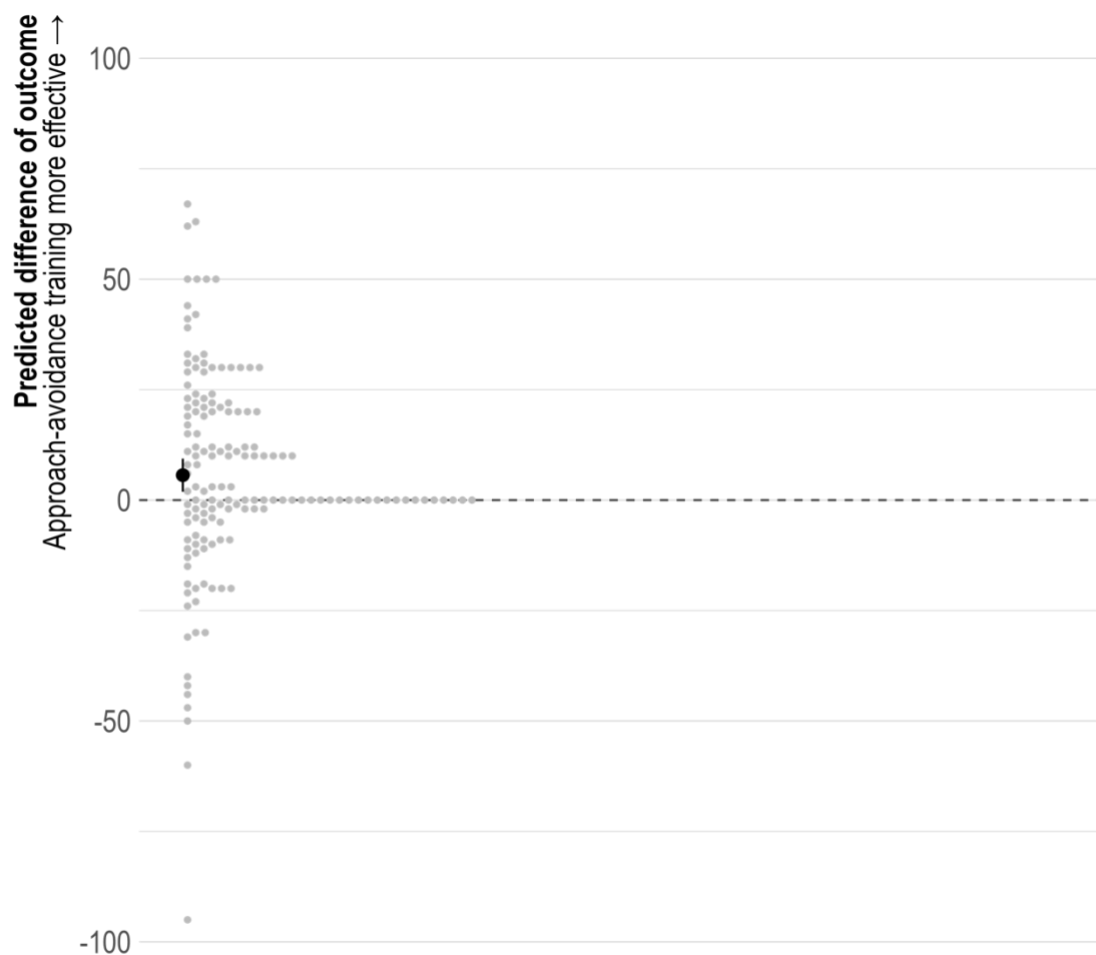
Results and Discussion

As preregistered, we conducted a paired-sampled t -test to test whether participants considered approach-avoidance training more effective than approach-avoidance instructions at forming implicit attitudes. Our dependent variable was the effectiveness of the intervention estimated by the participant (which we recoded so it ranged from -50 to +50, with 0 indicating that the intervention had no effect on participants' implicit attitude; see Figure 10.2) and our independent variable was the type of intervention participants had to evaluate. We did not exclude any participant from this analysis and we did not deviate from the preregistration. One hundred and fifty participants were included in the analysis.

The analysis revealed a significant effect of the type of intervention on the predicted effectiveness to form implicit attitudes (see Figure 10.2). Participants believed that the approach-avoidance training ($M = 16.4$, $SD = 20.9$) was more effective than approach-avoidance instructions ($M = 10.8$, $SD = 18.9$), $t(149) = 2.98$, $p = .003$, $d_z = 0.28$, 95% CI [0.06; 0.51]. On average, participants believed that the approach-avoidance training was 5.6 points more effective than the approach-avoidance instructions, 95% CI [1.9; 9.4].

Figure 10.2

Difference of Outcomes Estimated by Participants Between an Approach-Avoidance Training and an Approach-Avoidance Instruction Intervention (Experiment 8)



Note. Positive scores indicate that approach-avoidance training is estimated as more effective than approach-avoidance instructions. Each point represents a participant. Error bar represents the mean and the 95% confidence interval.

This result confirms the hypothesis that people infer that approach-avoidance training interventions are more effective at forming indirect evaluative responses than approach-avoidance instructions intervention. Still, it is easy to make the hypothesis that the wording in the descriptions of the interventions affected our results in an unintended way. Because we did not implement any check assessing participants' comprehension of the procedure, a reasonable argument against the validity of our results could be that our descriptions of the interventions mischaracterized approach-avoidance training or approach-avoidance instructions. In Experiment 4, we fix this issue by making participants experience an approach and avoidance training before evaluating how they would react in this experiment.

Experiment 9

To make sure that the results of Experiment 8 was not the result of participants' miscomprehension of the procedures, we decided to replicate this study with a more solid basis for participants' evaluation of the interventions. Experiment 4 uses the same protocol as Experiment 3 but includes an approach-avoidance training demo at the beginning of the experiment. If anything, experiencing an approach-avoidance training should improve the overall comprehension of the interventions. This experiment was preregistered. We report all measures, manipulations, and exclusions. Material, data, the analysis script, and the preregistration can be found at <https://osf.io/cf27h/>.³⁸

³⁸ This project might be subject to an embargo. If it is unavailable, please e-mail cedric.batailler@gmail.com in order to gain access to this project.

Method

Participant, Design, and Sensitivity Analysis

As in Experiment 3, our 151 participants ($M_{age} = 35.3$, $SD_{age} = 13.2$; 83 females, 65 males) were recruited via the Prolific platform (Palan & Schitter, 2018). Participants had to be native English speaker, have an approval rate over .90 on the Prolific platform, and have never participated in a study from our lab on the Prolific platform.

Participants engaged in a study with one within-participant variable (the type of training participants had to evaluate). The group that had to be approached during the approach-avoidance training demo (Luupite vs. Niffite) and the effectiveness evaluation order (approach-avoidance training first vs. approach-avoidance instructions first) were counterbalanced across participants.

We preregistered a sample size of 150 participants for the current experiment. Experiment 3 revealed that the difference of estimated effectiveness between approach-avoidance training and approach-avoidance instructions was a small-to-medium $d_z = 0.28$, 95% CI [0.06; 0.51]. Sensitivity analysis revealed that a paired-sample two-tailed t -test with an alpha criterion of .05 and with a sample size of 150 provides .79 statistical power to detect a small effect size (i.e., $d_z = 0.20$), .92 statistical power to detect an effect the same size of Experiment 3 (i.e., $d_z = 0.28$), over .99 statistical power to detect a medium effect size (i.e., $d_z = 0.50$), and over .99 statistical power to detect a large effect (i.e., $d_z = 0.80$). Given the analysis, we consider that this analysis is adequately powered.

Procedure and Material

As in Experiment 8, participants were recruited from their web browser thanks to the Prolific platform (Palan & Schitter, 2018). Participants were redirected to a webpage where our experiment was hosted—unlike Experiment 8, because we included an approach-

avoidance training demo, we used jsPsych rather than Qualtrics to code the experiment (see Experiment 1b; de Leeuw, 2015). Participants had to make predictions about the outcome of two interventions involving “a special paradigm called approach-avoidance training”. To make the predictions easier, before anything else, participants first engaged with an approach-avoidance training.

During about three minutes, participants engaged with a VAAST-based approach and avoidance training. In this training, they had to approach names from people belonging to a social group, the Luupites, and avoid names from people belonging to another group, the Niffites (or the opposite). To do so, participants had to use the arrow keys on their keyboard. The procedure was identical to the one we used when validating the online version of the VAAST in Experiment 1b, except for the numbers of trials. For each trial, a fixation cross appeared on the screen and was quickly followed by a Luupite or Niffite name. When the name appeared, participants had to make either an approach or an avoidance movement which would then trigger an approach or avoidance movement. The demo lasted for 30 trials.

Once the demo over, participants moved to the main task, that is, the prediction of interventions’ outcome. As in Experiment 8, the order of evaluations was counterbalanced across participants. Participants either read first about approach-avoidance training or about approach-avoidance training. The verbatim for Experiment 9 was almost identical to the one of Experiment 8, except for explicit references to the approach-avoidance training demonstration with which participants engage. Contrary to Experiment 3, the verbatim mentioned groups A and B rather than Luupite and Niffite (which were already mentioned during the approach-avoidance training demonstration).

After reading the description of the interventions, we included an attention check as a way to ensure that our participants paid attention during the experiment (participants were informed that the experiment could include attention checks). This check was not included in

Experiment 3. Following on Prolific's policy regarding attention check, we included a question asking participants: "based on the text above, what is your favorite color?". The text above the question asked them to type the word "baguette" when asked for their favorite color.

Right after the attention check, participants made predictions regarding the outcomes of the interventions with a visual analog scale (Figure 10.1, p. 215). As in Experiment 3, the order of the questions matched the order of the intervention descriptions. Participants were then thanked for their participation, debriefed, and paid. The whole study lasted approximately 10 minutes.

Results and Discussion

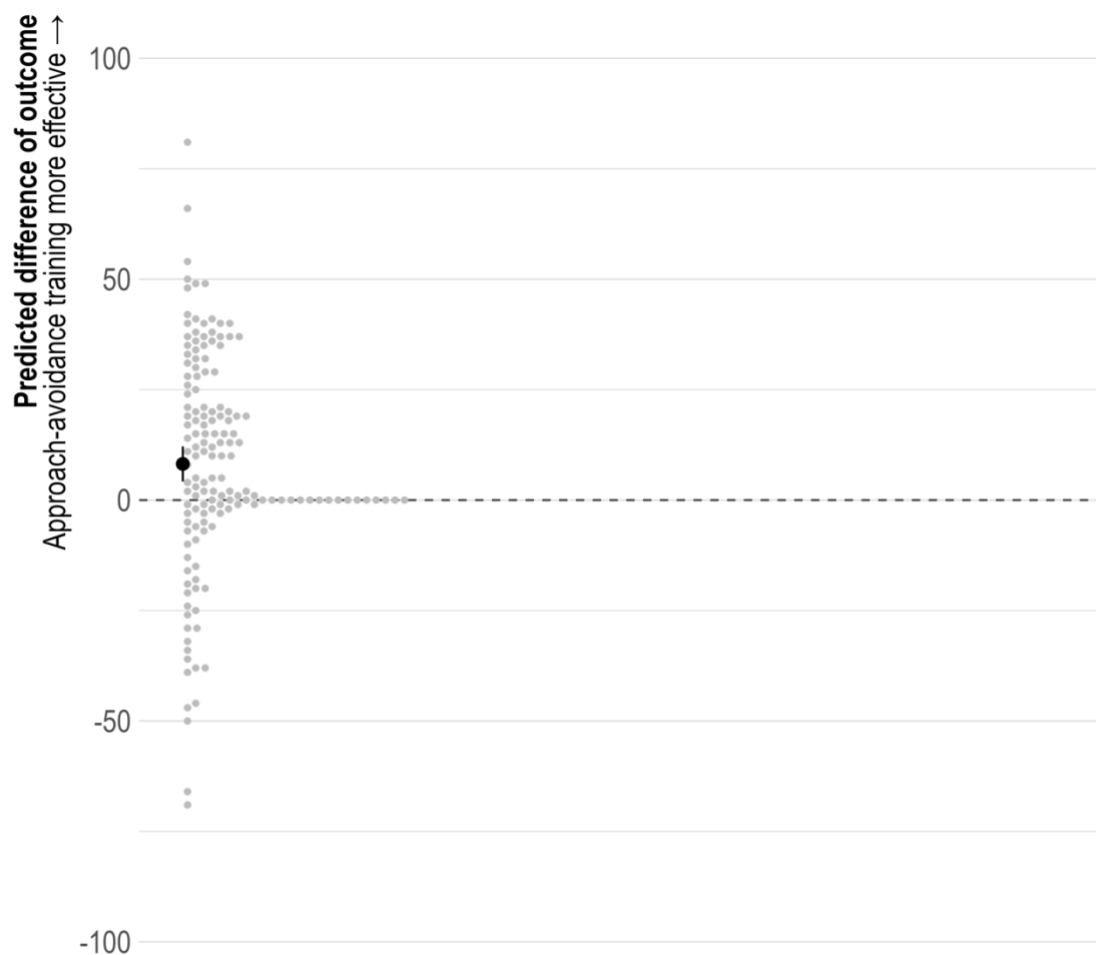
As preregistered, we conducted a paired-sampled t -test to test the hypothesis that participants consider approach-avoidance training more effective than approach-avoidance instructions. Our dependent variable was the effectiveness estimated by the participant (ranging from -50 to +50) and our independent variable was the type of intervention participants had to evaluate. We planned to exclude participants who failed our attention check and to exclude statistical outliers based on Judd et al.'s recommendation (2017). These criteria did not result in the exclusion of participants; 151 participants were included in the analysis.

This analysis revealed a significant effect of the type of intervention on effectiveness to form implicit attitudes (see Figure 10.3). Participants believed that the approach-avoidance training was more effective ($M = 19.5$, $SD = 20.0$) than the approach-avoidance training ($M = 11.3$, $SD = 17.8$), $t(150) = 4.07$, $p < .001$, $d_z = 0.43$, 95% CI [0.20; 0.66]. On average, participants believed that the approach-avoidance training was 8.18 points more effective than the approach-avoidance instructions, 95% CI [4.21; 12.16].

In the end, we found the same difference of expectation regarding the outcome of approach-avoidance training and approach-avoidance instructions as in Experiment 8. Even when participants used a demo of approach-avoidance training to make up their mind regarding the effectiveness of the interventions, they assumed that approach-avoidance training would have a greater impact on their implicit attitudes.

Figure 10.3

Difference of Effectiveness Estimated by Participants Between an Approach-Avoidance Training and an Approach-Avoidance Instruction Intervention (Experiment 9)

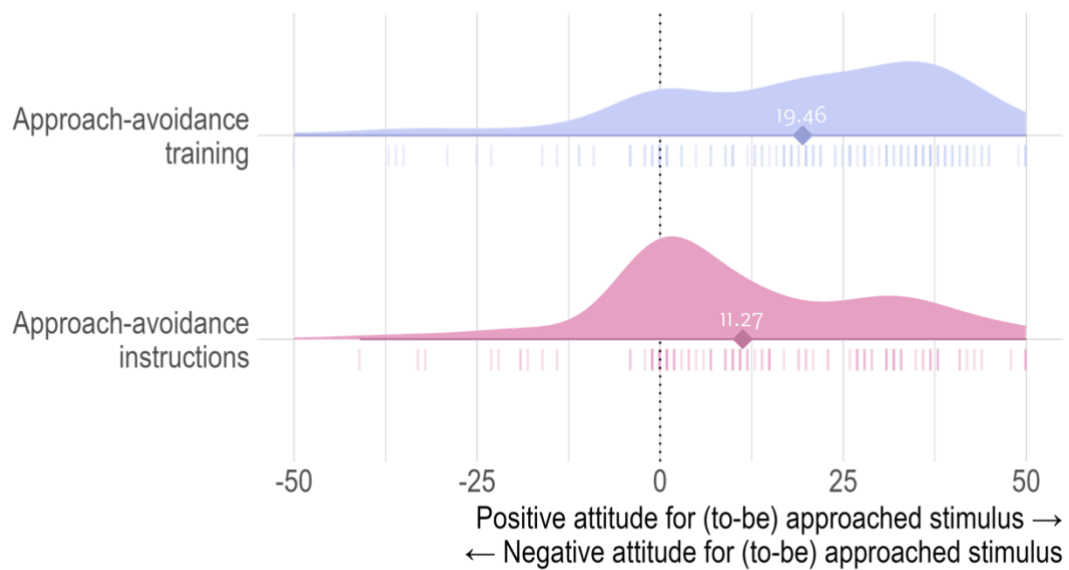


Note. Positive scores indicate that approach-avoidance training is estimated as more effective than approach-avoidance instructions. Each point represents a participant. Error bar represents the mean and the 95% confidence interval.

As an additional analysis, we computed how much more effective participants thought approach-avoidance training intervention was compared to approach-avoidance instruction intervention. This analysis was not pre-registered. To do so, we used participants' response distribution in Experiment 9 (see Figure 10.4). Because participants made a point estimate of an intervention's effectiveness, and because it was possible for participants to declare that an intervention would not be effective, it is possible to estimate how much more effective an intervention is compared to the other. This analysis reveals that, on average, participants believed that the approach-avoidance training intervention was 73% more effective than the approach-avoidance instructions intervention at changing their implicit attitudes. This effect seems to be partly driven by a large number of participants assuming that mere approach-avoidance instructions would not impact their implicit attitude (see Figure 10.4).

Figure 10.4

Distribution of the Estimated Outcomes for Approach-Avoidance Training and Approach-Avoidance Instructions in Experiment 9



Note. Ribbons represent the density of participants estimated outcome for the interventions. The larger the ribbon, the more participant estimated a specific outcome. Each bar represents a single participant of Experiment 9. The diamond indicates the average estimation for the intervention. Negative numbers on the x-axis indicate that the intervention would create negative implicit attitude for the (to be) approached stimulus and implicit positive attitude for the (to be) avoided stimulus, positive numbers indicate the opposite, and the dotted line indicates that the intervention has no effect.

General Discussion

In this chapter, we wanted to investigate people's lay theory regarding the outcome of approach-avoidance training and approach-avoidance instructions. To do so, we conducted two pre-registered experiments where we introduced participants to these procedures. In Experiment 9, participants even had the opportunity to take part in an approach-avoidance training. In the present dissertation, we focus on the effects of these procedures on indirect evaluative responses. Accordingly, our participants had to predict what would be the result of these procedures on an attitude IAT (Greenwald et al., 1998). First, Experiments 8 and 9 both revealed that, on average, participants considered the two procedures as effective at making people acquire an indirect evaluative response. Second, in both Experiments 8 and 9, participants predicted that an intervention based on approach-avoidance training would outperform an intervention based on approach-avoidance instructions. This effect was small-to-medium (Cohen, 1988). Participants' lay beliefs are in line with implicit social cognition theories. There is something special with approach-avoidance training when it comes to indirect evaluative response acquisition.

In the introduction section of this dissertation, we saw that one of the reasons why theoretical models of evaluation assumed that repeated experience is a critical factor for acquisition was the nature of indirect evaluative responses. When their use emerged, the theories describing the mechanisms underlying indirect evaluative responses assumed that these measures captured something progressively encoded over time which was now deeply rooted in the human mind. Greenwald and Banaji (1995) assumed that individuals themselves were not able to report where these behaviors came from. Fazio et al. (1986) assumed that they were so deeply rooted that the mere presentation of a stimulus would lead individuals to emit an uncontrollable behavior (e.g., Chen & Bargh, 1999; see also Bargh, 2017).

Because of the origin of implicit social cognition, modern theories assume that repetition-based experimental paradigms possess something special when it comes to indirect evaluation acquisition. Even if this “something special” differs in importance according to the theoretical models—the SEM assuming that it is a *sine qua non* condition (McConnell & Rydell, 2014), the single process propositional assuming that it helps indirect evaluation acquisition (Van Dessel et al., 2018a)—it is always described. As a consequence, paradigms such as approach-avoidance training are supposed to outperform paradigms such as approach-avoidance instructions. Experiments 8 and 9 show that naive participants, unaware of the history of indirect evaluative responses, make the same predictions. On top of that, they think that approach-avoidance training should outperform approach-avoidance instructions when it comes to the acquisition of indirect evaluative responses.

A limit one could highlight is our use of *implicit attitude* in Experiments 8 and 9. Many researchers in implicit social cognition agree to consider that it is very unlikely that any of the measures traditionally used (e.g., IATs, AMPs) is really “process-pure” (J. W. Sherman, 2008). In Experiments 8 and 9, we defined IAT as capturing an affective spontaneous reaction which is, with regard to our current knowledge about the IAT, partly incorrect (Corneille & Hütter, 2020). A reasonable argument regarding Experiments 8 and 9 could be that our description of the IAT scores influenced the answers of participants in a way that could have been different if we provided a more valid definition of IAT. However, empirical data regarding how people predict their IAT scores suggest that the theoretical perspective used to describe the IAT usually makes no difference (Rivers & Hahn, 2019). In other words, describing implicit attitudes in terms of “true underlying attitudes” (sic), “culturally learned associations”, or “sometimes [differing] from what people think of themselves” does not significantly impact one’s predictions. These data from the literature

suggest that the prediction of our participants would have been the same with different definition of implicit attitudes.

Conclusion

To sum up, in this chapter, across two experiments, we investigated whether naive participants would make the same predictions as theoretical models of evaluation. We assumed that participants would consider intervention such as the approach-avoidance training more effective than approach-avoidance instructions, exactly like the SEM, the APE model, or single-process propositional model of evaluation. The results confirmed our hypothesis. The remaining question is now to test whether it is really the case.

PART III: GENERAL DISCUSSION

Chapter 11

General Discussion

Modern theories of evaluation, such as the SEM (McConnell & Rydell, 2014) or the APE model (Gawronski & Bodenhausen, 2007), hold that the repetition of an experience contributes to the acquisition of indirect evaluative responses. Recent evidence, however, revealed that paradigms inspired from the repetition-based paradigms, but which do not implement repeated experience, can robustly form indirect evaluative responses. Instructions-based paradigms merely contain the instructions usually found in repetition-based paradigms and have been shown to influence indirect evaluative response (e.g., Gast & De Houwer, 2012; Van Dessel et al., 2015). Because these instructions are enough to form indirect evaluative response, one can wonder what the contribution of repeated experience is in repetition-based paradigms. Over mere verbal statements, does the repeated experience included in a paradigm like approach-avoidance training contribute to the acquisition of evaluative responses? This is the question we addressed in the present dissertation.

Summary of the Results

As reviewed in the introduction section of this dissertation, every modern model of evaluation (be it dual process or single process) predicts a specific contribution for repeated experience in evaluative learning. Based on these models, we argued that the indirect evaluative responses acquired as a result of an approach-avoidance training should differ. Specifically, these theoretical models led to the prediction that approach-avoidance training should have a greater effect on evaluation than mere approach-avoidance instructions. We tested this hypothesis in the present dissertation.

We used **Chapter 6** as an opportunity to develop an approach-avoidance training based on the recent advances in approach-avoidance implementation. Using a paradigm inspired by the VAAST (Rougier et al., 2018; see also Rougier et al., 2020), we tested whether approaching and avoiding novel stimuli resulted in the formation of an indirect evaluative response (on an identification IAT; Nosek et al., 2002). We first benchmarked this training against an approach-avoidance procedure used in the literature (i.e., the Joystick Task; Experiment 1a) and then tested whether the effects of the procedure we developed could be found again in an online experiment (Experiment 1b). These two experiments revealed that our procedure was successful at forming indirect evaluative responses. Compared to the indirect evaluative responses formed as a result of the Joystick task, those formed in the VAAST-based approach-avoidance training were descriptively larger (Experiment 1a). This procedure was also effective with participants who took the experiment online (Experiment 1b), showing that its effects were not limited to lab samples. We concluded this chapter by assessing whether this VAAST based approach-avoidance procedure could be used to replicate findings from the literature. In Experiment 2, we successfully replicated the work of Kawakami et al. (2008) showing that a math approach training resulted in a change of indirect evaluative response toward math for women who reported not to like math. Altogether, the experiments from **Chapter 6** support the idea that the approach-avoidance training that we developed can be used to investigate the contribution of repeated experience to indirect evaluative learning.

We compared the effects of approach-avoidance training and approach-avoidance instructions for the first time in **Chapter 7**. In a paradigm where participants were presented several logos, we compared the effects of the repeated experience of approach and avoidance movements to those of mere approach-avoidance instructions on an AMP. In Experiment 3, we first made sure that our procedures yield results consistent with the literature; we tested

whether both approach-avoidance training and approach-avoidance instructions resulted in the formation of an indirect evaluative response. The results of this experiment showed that the two procedures successfully resulted in the formation of an evaluative response. However, the sample size that we recruited in Experiment 3 did not allow to reliably assess the difference between the two (any difference smaller than $d_s = 0.71$ could not be reliably assessed given our sample size). Consistently, we did not observe a difference between the two procedures in terms of evaluative learning. The statistical power recruited in comparing the two procedures was larger in Experiment 4. Contrary to Experiment 3, this experiment adopted a within-subject design to compare the effects of the repeated experience of approach-avoidance movements to the effects of mere approach-avoidance instruction. We also used a sample size larger than the one used in Experiment 3 ($n = 104$ vs. $n = 63$). Despite the gain in statistical power in Experiment 4, we did not observe a significant difference between the evaluation of the stimuli which had been approached and avoided during the training and the stimuli which had not. Still, we again found a significant difference between the evaluation of the (to be) approached and those of the (to be) avoided stimuli.

In **Chapter 8**, we used an incidental design to our approach-avoidance training. In Experiment 5, we gave participants approach-avoidance instructions unrelated to the features of the stimuli which we targeted with the approach-avoidance training. In such a context, the contribution of repeated experience was assumed critical by some of the theoretical models of evaluation (i.e., single-process propositional model; Van Dessel et al., 2018a). Results of this experiment showed that, while participants acquired an evaluative response toward the features of stimuli that they were instructed to approach and avoid (the features instrumental for the training), we found no evidence that they acquired an evaluative response toward the features incidental to the training. Regarding the effect of the approach-avoidance procedures

on the evaluation of stimuli with feature instrumental for the training, we found no difference between approach-avoidance training and approach-avoidance instructions.

In **Chapter 9**, we compared how the effect of approach-avoidance training and approach-avoidance instructions evolved in time. In Experiment 6, we recruited participants who were then assigned to either an approach-avoidance training condition or an approach-avoidance instructions condition. Participants took an identification IAT right after the approach-avoidance procedure and were invited to come back the day after to take a second identification IAT. Based on De Houwer's (2018) single process propositional model, we made the hypothesis that the difference between the procedures should be easier to observe after a delay. We did not find support for this hypothesis, but we found a significant difference between approach-avoidance training and approach-avoidance instructions (in the first and the second IAT). In Experiment 6, the indirect evaluative responses acquired through an approach-avoidance training were more pronounced than the evaluative responses acquired through approach-avoidance instructions. This result was not significantly affected by the day participants took the IAT. We replicated the difference between the two procedures observed in the first stage of Experiment 6 in a highly powered Experiment 7.

Finally, in **Chapter 10**, participants naive to the challenges of implicit social cognition made predictions about the effects of approach-avoidance procedures. After being explained that researchers could use an approach-avoidance training or approach-avoidance instructions to form indirect evaluative responses (as measured with an IAT), they reported how much they think these two procedures would be effective. No matter whether participants took a demonstration of an approach-avoidance training (Experiment 9) or not (Experiment 8), they predicted that both procedures should result in the acquisition of an indirect evaluative response, and that approach-avoidance training should be more effective than approach-avoidance instructions.

Altogether, we believe that the results of these experiments provide partial support for the hypothesis of a contribution of repeated experience in indirect evaluative learning. We consider that the contribution of the present dissertation, however, goes beyond a mere functional relationship between a type of procedure and indirect evaluative responses. In the next section, we will contextualize our results within the implicit social cognition literature.

Implications of our Results

Models of Evaluation

In this dissertation, we used three different models of evaluation as theoretical foundation. These models allowed us to design experiments to investigate the role of repeated experience in evaluative learning. In Experiments 3–7, we examined what were the predictions (if any) of these models regarding the role of repeated experience. We consider the results of these experiments were informative and it is now time to summarize how they fit with the implicit social cognition literature.

The first model that we reviewed, the SEM, was also the most radically dual (McConnell & Rydell, 2014). In the introduction section, we discussed how this model considered associative processes as responsible for indirect evaluative responses and propositional processes as responsible for direct evaluative responses. In the SEM, each of these processes operate in their own system. As factors processed in the propositional system are not supposed to influence indirect evaluative responses, the SEM is not fit to account for the effect of instructions on evaluative learning of indirect responses. Yet, in Experiments 3–7, we consistently found that mere approach-avoidance instructions formed indirect evaluative responses. These results replicated and extended other lines of research in the literature (e.g., Van Dessel et al., 2015) by showing that the effects of approach-avoidance instructions are not limited to the IAT but generalize to the AMP (Experiments 3–5). These findings, and their implication regarding the SEM, align with recent a recent review of the

literature suggesting that the radicalness endorsed in the dissociation between associative and propositional processes constraints this model in a way which is not consistent with evidence literature (Corneille & Mertens, 2020; Corneille & Stahl, 2019).

The APE model is a less radical variant of the dual model of evaluation. It still accounts for direct and indirect evaluative response using associative and propositional processes, but it assumes that the associative processes can impact the propositional ones and that the propositional processes can impact the associative ones (Gawronski & Bodenhausen, 2006, 2007). This flexibility allows for the APE model to account for indirect evaluative learning through mere instructions. For the APE model, approach-avoidance instructions, once processed, can result in the formation of new mere associations in memory, new associations which will later impact the evaluation of stimuli on indirect evaluative measures. While data regarding the effects of approach-avoidance instructions are easier to reconcile with the APE model than with the SEM (Experiments 3–7), the APE model still makes predictions for which we did not find support.

Along with mechanisms specific to associative processes (e.g., activation spreading), the APE model describes the conditions in which these processes are assumed to operate (i.e., operating conditions). Notably, the APE model assumes that the learning of an indirect evaluative response can occur thanks to a process which is efficient and without the goal to learn something (Gawronski & Bodenhausen, 2014a). If true, we should have been able to observe the acquisition of an indirect evaluative response for stimuli which possessed features incidental to the training in Experiment 5. While the statistical procedures that we used do not allow us to affirm that there is no effect of approach-avoidance training on the acquisition of an indirect evaluative response for stimuli which possessed features incidental to the training (Lakens et al., 2018), the acquired evaluative responses, if any, are significantly smaller than the ones acquired for stimuli possessing features instrumental for

the training. This observation is not easily made by the APE model and would require additional assumptions. The same is not true for single-process propositional models of evaluation which can easily account for this result.

Single-process propositional models such as the one described by De Houwer (2018) or Van Dessel et al. (2018a) suggest that propositional processes are responsible for indirect evaluative response acquisition (as well as direct evaluative response acquisition). These models assume how propositional processes operate (i.e., the operating principles), but also when these processes will operate (i.e., the operating conditions). By contrast to the dual models of evaluation (the SEM, the APE model), single-process propositional models assume no efficient way to acquire an indirect evaluative response. To form a proposition, which will later result in the acquisition of an evaluative response, one needs to recruit attention (De Houwer, 2018; Van Dessel et al., 2018a). In Experiment 5, as long as we make the hypothesis that participants did not pay attention to the features incidental to the training, we should expect a stronger evaluative learning for stimuli with features instrumental for the training.

Single-process propositional processes propositional models can account for the effects of mere instructions, for the differences between approach-avoidance training and approach-avoidance instructions, and for the difference of evaluative learning between stimuli with features incidental or instrumental for a training. However, it does not necessarily mean that this kind of model is superior. Tenants of single-process propositional models acknowledge that, in their current implementation, these models might be harder to falsify than their dual-process counterparts. The fact that our results fit better with single-process propositional models of evaluation could be the result of a model not falsifiable enough rather than the result of a model fitting accurately with our results. While promising, this class of model should be evaluated on its capacity to account for existing data, to make new prediction, and on its falsifiability (Gawronski & Bodenhausen, 2015). Assessing the

falsifiability of single-process propositional models goes beyond the goals of the present manuscript. We will note, however, that the single-process propositional models of evaluation allowed us to make new predictions regarding the situations in which we should have observed a larger difference between approach-avoidance training and approach-avoidance instructions (Experiments 5 & 6)—even if the data regarding these predictions remained inconclusive.

To sum up, as we saw in the current section, our lines of research provide interesting insights regarding the theories of evaluations. Overall, it seems that the single process-propositional model of evaluation is the better suited to account for our results and the SEM, the worst suited—with the APE model sitting in between. An aspect of our results, however, is not discussed by any of the models of evaluation: the size of the difference between repetition-based and instructions-based procedures.

On the Effect Size of the Difference Between Approach-Avoidance Procedure

In several papers, researchers argued that merely reporting p -values is not the most informative way to assess the relevance of a finding (e.g., Cumming, 2008; Lakens, 2013). Instead, researchers recommend a greater emphasis on effect sizes. In the context of the difference between approach-avoidance training and approach-avoidance instructions, focusing on effect size seems especially relevant, because it provides interesting insights on the heterogeneous results that we observed in Experiments 3–7. While Experiments 3–5 did not show a significant difference between training- and instructions-based procedures, Experiments 6 and 7 did. Statistics are sufficient to account for these discrepancies.

Across Experiments 6 and 7, the effect size of the difference between the approach-avoidance training and approach-avoidance instruction conditions was small, $d_s = 0.25$,

CI95% [0.14; 0.36] (Cohen, 1988).³⁹ As briefly discussed in **Chapter 9**, an effect size this small requires a large sample size to be reliably observed. More precisely, one would need to recruit 694 participants to obtain a statistical power of .95 in a between-subject design or 398 participants to obtain a statistical power of .80. The sample sizes recruited in Experiments 3 and 5 are far below this value (respectively, $n = 63$ and $n = 238$).⁴⁰ While there is room for theoretical explanations of the inconsistencies of our results in Experiments 3–5 and Experiments 6 and 8,⁴¹ we must stress that the Experiments 3–5 were underpowered for an effect size of $d_s = 0.25$.

Focusing on effect size rather than mere p-values also helps us understand some of the implications of our findings. If we interpret the effect size of the difference between approach-avoidance training and approach-avoidance instruction ($d_s = 0.25$) in more detail, it indicates that 59.9% of the participants who actually performed the approach-avoidance

³⁹ We conducted a random-effects meta-analysis with the *metafor* package (REML estimation; Viechtbauer, 2010) to estimate the difference between approach-avoidance training and instructions across Experiments 6 and 7. Besides the average effect size, this analysis revealed no evidence for heterogeneity between the two experiments, $Q(1) < .001$, $p = .999$.

⁴⁰ We focus on these two experiments and not on Experiment 4 because Experiment 4 adopts a within-subject design.

⁴¹ Note that we did not actually compare the difference between Experiments 3–5 and Experiments 6 and 7. We did not run such an analysis because we did not adopt the same analytical strategy to analyze the data from the two kinds of experiments. Comparing Experiments 3–5 to Experiments 6 and 7 would be an analysis with a lot of degrees of freedom (Simmons et al., 2011).

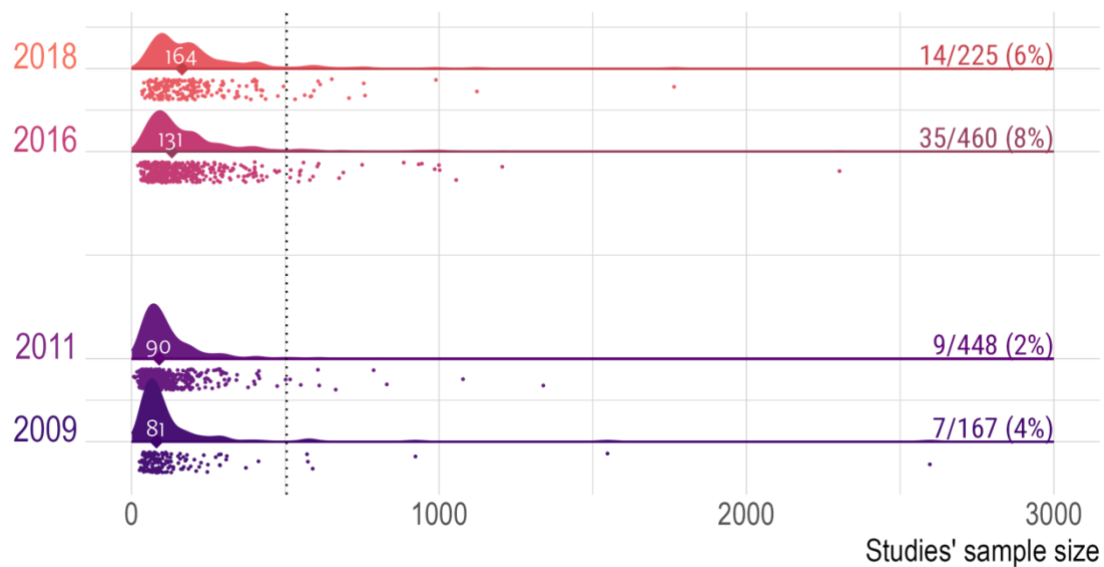
training show a stronger acquisition of evaluative responses on the IAT than the average participant who received approach-avoidance instructions (Cohen's U_3). If we had to pick one person at random in the approach-avoidance training and approach-avoidance instructions conditions, there would be a 57.0% probability that the person in the approach-avoidance training condition showed stronger evaluative response toward the (to be) approach and the (to be) avoided than the person in the approach-avoidance instructions condition (Magnusson, 2021). In the end, several sets of rules exist to describe the order of magnitude of an effect size, and they all suggest that the difference between the difference of the two procedures is small (Cohen, 1988; Gignac & Szodorai, 2016; Lovakov & Agadullina, 2021; Sawilowsky, 2009). Effect sizes in social psychology, however, do not come with operational decisions associated with the different labels (Fidler, 2002). Phrased differently, at the moment, we do not have clear guidelines on how different effect sizes should map onto theories. What is the required effect size for researcher to account for a type of effect with a new process? What is the required effect size for a theory to be disproven? We consider this question critical for the pursuit of the work focusing on the contribution of repeated experience.

A fair question to ask would be related to the practical significance of our results. Is the effect big enough to have some importance? A first way to answer this question would be to contextualize this effect size in the current practices in social psychology. As a matter of comparison, thanks to a review conducted by Sassenberg and Ditrich (2019), it is possible to investigate how many studies published in top social psychology journals (*Journal of Experimental Social Psychology*, *Journal of Personality and Social Psychology*, *Personality and Social Psychology Bulletin*, and *Social Psychology and Personality Science*) would have had .80 power to detect the effect observed in Experiments 6 and 7. While one can argue that such analysis is suboptimal, as scientists choose their sample size a priori (and therefore

sometimes have good reasons not to collect data from large sample sizes), we consider that it still provides some context to interpret the effect of repeated experience. In a nutshell, less than 10% of the studies reported in those journals had enough power to detect our effects (see Figure 11.1).

Figure 11.1

Reported Sample Size in the Four Top Empirical Social Psychology Journals



Note. Each point represents the sample size for a study reported in a the *Journal of Experimental Social Psychology*, the *Journal of Personality and Social Psychology*, the *Personality and Social Psychology Bulletin*, or *Social Psychology and Personality Science*. Ribbons represents the density of sample sizes for a given year. The larger the ribbon, the more studies report a given sample size. Dotted line represents the sample size necessary to achieve 80% statistical power for an effect size of $d_s = 0.25$. For each horizontal line, the proportion of studies to achieve such threshold is indicated. These data were collected by Sassenberg and Ditrich (2019).

Even though the effect size of the difference between approach-avoidance procedures can easily be considered as small, this tells us little about how this result should be integrated in theories. In other words, how essential is it that theoretical models of evaluation account for a difference of this effect size? Because, virtually, given a sample size large enough, any effect can be considered as different from a null effect (see Cohen, 1992), it is important to define the relevant criteria regarding the theoretical importance of an effect. A promising approach uses the concept of smallest effect size of interest (SESOI). A SESOI-based approach involves scientists to decide on an a priori threshold which would make them consider an effect size relevant. For example, researchers could consider that, regarding the effect they are investigating, $d_s = 0.30$ is their SESOI. Any results which is significantly smaller than the absolute value of this SESOI would therefore be considered irrelevant. As an example, in this manuscript, we adopted this approach to benchmark the effectiveness of the VAAST-based approach-avoidance training against a Joystick task-based one in Experiment 1a. At the time, we adopted a SESOI based on reasons contingent on our resources, but others can be considered, such as the clinical relevance (Lakens, 2017; see also Anvari & Lakens, 2021). We consider that further studies investigating the contribution of repeated experience should consider this approach as it would make integrating the results of empirical research easier to integrate to the existing literature.

As discussed above in the manuscript, approach-avoidance training procedures have been used in several domains. Among other, we can find math (Kawakami et al., 2008), eating behaviors (Ferentzi et al., 2018), sexual behaviors (A. M. Hahn et al., 2019), or depression (Becker et al., 2019). Given that we found that the difference between approach-avoidance training and approach-avoidance instructions was likely small (Experiments 6 & 7), an important question is related to the gain in implementing approach-

avoidance training procedure compared to its costs. We believe that it is safer for researchers to use approach-avoidance training procedures than approach-avoidance instructions. First of all, a small effect size in the lab does not necessarily translate into a small effect in real life (an illustration of this argument applied to IAT research can be found in Greenwald et al., 2015). Moreover, our own data suggest that participants believe that approach-avoidance trainings are more effective than approach-avoidance instructions (Experiments 8 & 9). Given the importance of outcome expectancies in psychotherapy outcomes (Greenberg et al., 2006), we believe that using approach-avoidance training over approach-avoidance instructions is safer. We think that the overall evidence in the literature plays in favor of approach-avoidance training compared to approach-avoidance instructions.

Other Research Focusing on the Contribution of Repeated Experience to Evaluative Learning

As we mentioned in the introduction section of this dissertation, the advances of the last decade motivated our research question. With the accumulation of evidence showing the impact of instruction-based procedures on evaluative learning (e.g., Gast & De Houwer, 2012; Van Dessel et al., 2015), the question of the role of repeated experience became more and more critical to investigate. At the time we started this research project, in 2016, we knew of no other studies investigating the difference of effects between repetition-based and instructions-based.⁴² Since then, however, several studies did. In the next section, we will review these projects. In addition to discussing how these projects are related to the literature,

⁴² While we knew of no other research investigating the difference of effects between repetition-based and instructions-based, studies investigated the joint contribution of verbal statement and repetition-based paradigm to indirect evaluative learning (Peters & Gawronski, 2011; Zanon et al., 2014).

we will also examine how they fit with the line of research conducted in the present dissertation.

The Contribution of Experience in Evaluative Conditioning Paradigms

In the present dissertation, we focused on approach-avoidance procedures. One of the first lines of research comparing the differences between repetition-based and instructions-based paradigms, however, adopted procedures based evaluative conditioning. Across six experiments, Kurdi and Banaji (2017) compared three different procedures according to their effects on an IAT: a procedure where participants both received verbal statements indicating that categories of stimuli would be paired with positive and negative stimuli and where they actually experienced the pairing, a procedure where they merely received the verbal statements, and a procedure where participants experienced the pairing but did not receive any instructions about it.

In terms of design of the procedure, our comparisons between approach-avoidance training and approach-avoidance instructions procedures (Experiments 3–7) are similar to the comparison between the procedures containing both the verbal statements and the actual pairing and the one containing mere verbal statements. Our results, however, differ a bit from those obtained by Kurdi and Banaji (2017). While descriptively in the direction of a stronger effect of the procedure containing statements and pairing, the difference between the procedure containing statements and pairing and the one merely including statements was not significant. Our Experiments 6 and 7 revealed a significant difference between these two kinds of procedures. The results described in Kurdi and Banaji's paper could hardly be interpreted in terms of insufficient power, as they recruited a total of 6 492 participants across their six experiments. In the end, if anything, it is certain that the difference between the two procedures is not a difference of large magnitude. Statistical procedures such as non-inferiority tests would be informative regarding this question (Wellek, 2010).

If we focus, however, on the comparison between the procedures including both verbal statements and actual pairing and those merely including pairing, our results and those of Kurdi and Banaji (2017) align. In Experiment 5, we investigated whether the repetition of approach and avoidance movements resulted in 1) the acquisition of an evaluative response for stimuli possessing a feature related to the instructions received by participants (features instrumental for the training) and 2) the acquisition of an evaluative response for stimuli possessing a feature not related to the instructions received by the participants (features incidental to the training). Our results revealed that the evaluative responses for stimuli possessing features instrumental to the training were larger than those for stimuli possessing features incidental to the training (in fact, we did not find evidence that participants acquired an evaluative response toward those stimuli). For their part, Kurdi and Banaji (2017) also found that evaluative responses formed as a result of both verbal statement and actual pairing were larger in magnitude than those formed as a result of mere pairing. These results support a view of evaluation acquisition which is not efficient or which depends on participants' goal (see Corneille & Mertens, 2020; Corneille & Stahl, 2019). This view is the one adopted by single-process propositional model of evaluation (De Houwer, 2018; Van Dessel et al., 2018a).

Approach-Avoidance Experience and its Impact on Reverse Correlation Procedures

While other teams investigated the contribution of repeated experience with another paradigm (i.e., Evaluative Conditioning, Kurdi & Banaji, 2017), others focused on other evaluative responses. Rougier et al. (2021) compared approach-avoidance training and approach-avoidance instructions with a reverse correlation procedure (Dotsch & Todorov, 2012). This procedure allowed them to explore the effects of approach and avoidance movement on mental representation.

Understanding the idea of mental representation is easier with an example. When asked to picture the “a” letter in your mind, you could represent it as an uppercase or a lowercase letter, as a letter written with serif or without. When asked to picture an academic professor, you could picture them as a male or as a female, as someone dressed formally or not, and so on. In a nutshell, reverse correlation is an experimental procedure capturing this mental representation in an intelligible format (i.e., a picture on a computer; Dotsch & Todorov, 2012). The result of the procedure is usually called a classification image. In their paper, Rougier et al. (2021) decided to build the classification images for approached and avoided faces.

In most of their experiments, a first group of participants approached and avoided faces, then, they performed the reverse-correlation task, allowing the researchers to capture the mental representation of approach and avoided faces (see Figure 11.2). A second group of participants then evaluated the classification images—the faces captured—on several dimensions (e.g., trustworthiness, cleverness). Overall, the classification images of approached faces were rated more positively on a set of traits one could expect to be relevant regarding approach and avoidance behaviors (e.g., trustworthiness, generosity). Interestingly for the present dissertation, in one of their experiments, Rougier et al. (2021) compared the classification images generated by approach-avoidance trainings and approach-avoidance instructions. When comparing the two, it seems that the training resulted in classification images with more pronounced features. The difference observed between the image linked to approach and avoidance was bigger than the one linked to approach instruction and avoidance instructions.

Figure 11.2

Classification Images of Approached and Avoided Faces in Rougier et al. (2021)



Note. Figure adapted from Rougier et al. (2021).

These results are consistent with our own results showing that approach-avoidance training, compared to mere approach-avoidance instructions, generate more pronounced indirect evaluative response (Experiments 6 and 7). We consider Rougier et al.'s (2021) results consistent with ours regarding how they fit with the literature. Their results do support the idea of a contribution of repeated experience in evaluation acquisition. This also shows that the difference between the two types of procedures is not limited to the IAT. This possibility was suggested by the non-significant differences between the approach-avoidance training and approach-avoidance instructions procedures when we used the AMP (Experiments 3–5).

Approach-Avoidance Experience and its Impact on Impression Formation

The present dissertation was the opportunity to address the questions of the difference between approach-avoidance training and approach-avoidance instructions. While we think

that this question is especially relevant with regard to the indirect evaluative responses, we also investigated this difference on direct evaluative response in a side project. As we said, contrary to the experiments reported in the main experimental part of this dissertation, this line of research focused on direct rather than indirect evaluative responses. Specifically, we were interested in comparing the effect of the two approach-avoidance procedures on impression formation. Do approaching or avoiding individuals changes our evaluation of them? What about receiving approach and avoidance instructions? Direct evaluative responses allowed us to investigate participants' evaluations in a context where they could elaborate on their answer (Braud, 2020), a factor that could be relevant regarding the differences between the two procedures.

Elaboration describes how much someone will think about the ideas relevant to an issue to solve the issue (Petty & Cacioppo, 1986). For example, somebody spending two hours picking a restaurant—because they consider every possible option—shows greater elaboration than somebody who decided to go to the first opened place. Typically, indirect evaluative responses do not leave room for elaboration, as they often require participants to perform a task under a time constraint. Yet, we think that leaving room for elaboration could have its importance regarding the difference between approach-avoidance training and approach-avoidance instructions.

As discussed in our literature review, propositional processes are involved when participants must generate a direct evaluative response (De Houwer, 2018; Gawronski & Bodenhausen, 2007; McConnell & Rydell, 2014). Usually, someone will consider a set of propositions, and, based on the idea that they must maintain consistency between the different propositions, they pick an answer (see the Propositional Processes section). Judgeability describes the feeling of detaining information relevant for a given problem and was shown to have an effect on someone's direct evaluative responses. Compared to people

who don't, people thinking that they possess the relevant information to judge someone tend to emit more extreme judgment (Yzerbyt et al., 1994, 1998).

Judgeability could be relevant regarding a difference of effect between approach-avoidance training and approach-avoidance instructions. In terms of impression formation, approaching or avoiding someone could be considered as relevant information to evaluate the person. However, merely receiving instructions, especially if the impression formation measure allows participants to elaborate, could make participants realize that they do not possess the relevant information to evaluate the target. In two experiments ($N = 300$ and $N = 300$), we investigated the difference of effects between approach-avoidance training and approach-avoidance instructions on direct evaluative responses. In these experiments, participants received approach-avoidance instructions toward two social groups, and half of them actually performed the movements. Then, participants evaluated how representative of a group some traits were (e.g., warm, intelligent, unfriendly). While these experiments revealed that (to be) approached social groups were evaluated more positively than (to be) avoided social groups, we could not produce conclusive evidence regarding the difference of influence on impression formation between approach-avoidance training and approach-avoidance instructions. Interested readers will find additional information on the OSF repositories of the experiments: <https://osf.io/7be4u/> and <https://osf.io/p9zhx/>.

Approach-Avoidance Procedures and Evaluative Response Change

As discussed briefly in the discussion of Experiments 6 and 7, evaluative response acquisition refers to two distinct situations: evaluative response formation and evaluative response change. Both these situations described how new experiences with a stimulus impact its subsequent evaluations, but evaluative response formation assumes that no evaluative response was evoked by the stimulus before the new experiences and evaluative response assumes that specific evaluative responses were evoked before new experiences.

This distinction maps relatively well with the difference between a novel and a known stimulus. In the present dissertation, we used the Luupite and Niffite groups (Experiments 1a, 1b, & 6–7), pictures related to math and art (Experiment 2), and logos that we designed (Experiments 3 & 4) as stimuli. Each time, we investigated how new experiences (be they approach-avoidance movements or mere instructions) impacted the evaluations of these stimuli. Anytime we focused on the difference of acquisition between approach-avoidance training and approach-avoidance instructions procedures, we used novel stimuli (Experiments 3–7). Regarding this question, our results are restricted to indirect evaluative response formation, and one could wonder whether they generalize to indirect evaluative response change.

In a recent paper, Van Dessel et al. (2020) replicated and extended the work of Kawakami et al. (2007) on the evaluation of outgroup member. In their original paper, Kawakami et al. (2007) investigated whether approaching faces of Black American changed White Americans' indirect evaluative responses toward Black Americans (on an attitude IAT; Greenwald et al., 1998). Results revealed that approaching Black faces made it easier for participants to later perform an IAT when they used the same key to categorize Black faces and positive target words. Van Dessel et al. (2020) investigated whether they could replicate this effect and how it compared to an approach-avoidance instructions procedure (Experiment 4). Results of this experiment are consistent with ours. They show that approach-avoidance training outperformed approach-avoidance instructions when it comes to the change of indirect evaluative responses.⁴³

⁴³ Note that the difference of effectiveness between the two procedures, while significant, is a small size effect, $F(1, 793) = 4.02, p = .045, d_s = 0.14$. A consequence of this small size effect

The results of Experiments 8 and 9 and those from Van Dessel et al. (2020) complement each other in showing that the difference between approach-avoidance training and approach-avoidance instructions generalize across both evaluative response formation and evaluative response change. When it comes to the effectiveness of approach-avoidance instructions, our results (i.e., Experiments 8 and 9) already suggested that approach-avoidance training was better a better paradigm to acquire an indirect evaluative response. Van Dessel et al.'s (2020) results, however, go even further, showing that, when it comes to evaluative response change, approach-avoidance instructions are not an effective procedure (rather than a *less* effective procedure). Van Dessel et al.'s (2020) results provide evidence for an absence of effect of approach-avoidance instructions on indirect evaluative response change (see also Van Dessel et al., 2015).

As discussed above, it is possible for models like the APE model (Gawronski & Bodenhausen, 2006, 2007, 2014a) or the single process-propositional models of evaluation (De Houwer, 2018; Van Dessel et al., 2018a) to account for the difference between repeated-experience- and instructions-based procedure. These models can also explain why it is harder to change an evaluative response than to form one. However, while the differences between repeated-experience- and instructions-based procedures mainly involve mechanisms related to the encoding of information in memory, the differences between evaluative response formation and evaluative response change involve mechanisms related to how encoded information influences evaluation (Gast, 2018)—an aspect we discussed in less detail. In a

is that Bayesian analysis of the result do not perfectly corroborate the hypothesis of a difference between the procedure. This analysis suggests that the data provide slightly more evidence for the absence than for the presence of a difference in the effect of approach-avoidance training and that of approach-avoidance instructions, $BF_{10} = 0.67$.

nutshell, if we focus on the similarity-based retrieval mechanisms present in both the APE model and the single-process propositional models (De Houwer, 2018; Gawronski & Bodenhausen, 2006, 2014a), the context of evaluation formation should describe situations in which one type of trace should be retrieved during an evaluation, whereas the context of evaluation formation should describe situations in which many types of traces should be retrieved during an evaluation. In other words, the only traces retrieved when confronted to a Luupite should be the one where they were approached, whereas traces retrieved when confronted to a math-related stimulus should be more diverse, even with an approach-avoidance training. In the context of the difference between repeated-experience- and instructions-based procedure, another line of research is relevant to the question of the expression of evaluative response.

Disentangling the Processes Involved in Expression of an Evaluative Response Acquired through Experience

In the present dissertation, we decided to refer to the measures that emerged in the field of implicit social cognition as indirect evaluative responses. We considered this approach advantageous as it allowed us not to rely on erroneous theoretical assumptions regarding the processes underlying responses on an indirect evaluative response and to compare the predictions of several models in a common language (De Houwer et al., 2013). A different but common practice is to describe the indirect evaluative responses according to an assumption regarding how a response is generated in such tasks (e.g., these measures are produced automatically; for a critical review, see Corneille & Hütter, 2020). This general confusion is facilitated by the fact that the most common term used to describe these measures is the term implicit. Still, while adopting the term indirect evaluative response can be seen as an advance, both the use of indirect evaluative measures and implicit measures present a problem: They give the impression of a unitary construct, and this impression can

be seen at a within-measure level and at a between-measure level. First, at a within-measure level, adopting a single term for these kinds of measures can give the false impression that only one process operates when one generates an evaluative response. The evidence in the literature goes against this idea (J. W. Sherman, 2008).

Research in the implicit social cognition field supports the idea that indirect evaluative responses are not process-pure. In other words, these measures are not the result of only one process, but the combination of several processes which might differ in terms of operating condition and principles (J. W. Sherman, 2008). For example, the processes involved in the generation of responses during an IAT contain both automatic and non-automatic features (as in processes that operate unconditionally or not; Conrey et al., 2005). These theoretical considerations are supported by mathematical modeling of the responses on indirect measures which allows the disentangling of different features. This analytical strategy is referred to as Multinomial Processing Tree modeling of indirect evaluative measures have been used to disentangle processes underpinning evaluative response (MPT; for a review, see Hütter & Klauer, 2016). The use of MPT has allowed us to better understand some results in the literature. For example, it provided an explanation to the surprising results indicating that older adults show a response pattern on the IAT suggesting an antiaging bias (Gonsalkorale et al., 2014). This result is surprising because evidence in the literature is consistent in that the self is strongly linked with positive valence (Greenwald et al., 2002), and so that it should be expected that older adults demonstrate a stronger association with positive and the “old” concept. Using the quad model—an MPT disentangling different sources of influence in IAT measure (Conrey et al., 2005)—Gonsalkorale et al. (2014) found that, while the IAT score among younger adults reflected an automatic association between “old” and “bad”, older adults’ IAT score was reflecting a less strong association between “old” and “bad”, but older adults being less able to overcome this association. This result is

consistent with the effect of age on executive function (see Rabbitt, 1979) and would not have been observable without the MPT methodology.

In a recent line of research, Smith et al. (2019) applied the MPT methodology to the question of the contribution of repeated experience on evaluative learning on an IAT, going further than an analytical approach which would consider performances at an IAT as a unitary construct. To do so, they used the same quad model that we described in Gonsalkorale et al.'s (2014) line of research. Being more specific with this MPT model, the quad-model assumes four different parameters underlying the response in an IAT procedure: the Activation parameter (AC), the Detection parameter (D), the Overcoming Bias parameter (OB), and the Guessing parameter (G; Calanchini & Sherman, 2013; Conrey et al., 2005; J. W. Sherman, 2008). In the IAT procedure, the AC parameter has been validated as a tendency for a response to be elicited by the automatic association between mental concepts, the D parameter as the ability for the participant to correctly identify the correct response on the IAT trial, the OB parameter as the ability for a participant to bypass the response induced by automatic activation when it contradicts the response identified on the basis of Detection, and the G parameter as the result of anything guiding the response but the other processes. In their research, Smith et al. (2019) used the quad-model to compare the IAT scores resulting from procedures based approach-avoidance training, instructions of approach-avoidance, evaluative conditioning, instructions of evaluation conditioning, mere exposure (see Zajonc, 1968), and instructions of mere exposure (see Van Dessel, Mertens, et al., 2017)—mere exposure procedures are procedures where one can observe a change in evaluative responses to a stimulus (which becomes more positive) as a result of mere exposure to the stimulus (Bornstein, 1989; Zajonc, 1968).

The results of three experiments revealed that the different types of procedures impacted several parameters of the quad-model. Both the procedures based on repeated-

experience and those based on instructions influenced the AC parameters consistently. This result supports the idea that the change in IAT scores resulting from the procedure is caused (at least in part) by a factor related to the evaluation of stimuli (C. T. Smith et al., 2019), as opposed to something not related to the stimuli, such as executive functions. When it comes to the comparison between procedures based on repeated experience and the procedures based on instructions, results suggest that both affect the AC parameter to a similar extent.⁴⁴ However, the AC parameter is not the only one impacted by the evaluation acquisition procedures. Smith et al.'s data suggest that instructions, compared to repeated-experience, result in higher values for the D parameter. In other words, participants in the instructions-based procedure conditions were better at accurately detecting the correct category of the target word during the IAT. Smith et al. interpreted this result in terms of depletion induced by the repetition-based procedures (and not the instructions-based ones). Indeed, the D parameter has already been observed as being constrained by cognitive capacity (Conrey et al., 2005). Because repetition-based paradigms require participants to engage in a task, while instructions-based paradigms do not, it is possible that the former depleted participants resources.

We could wonder whether the difference between repeated-experience based paradigms and instructions-based paradigms regarding the D parameter could explain the

⁴⁴ Depending on the analyses conducted, differences between experience- and instructions-based procedures can be observed. Notably, when aggregating procedures (i.e., combining repetition-based approach-avoidance training, Evaluative Conditioning, and mere exposure together, and the same for instructions-based procedures), analysis reveals a larger impact of instructions-based procedure on the AC parameter. However, this difference does not hold when each repetition-based procedure is compared to its instructions-based counterpart.

differences between the two approach-avoidance procedures in Experiments 8 and 9. Indeed, in these experiments, we observed a difference between the two procedures based on an analysis of the IAT D scores, an index derived from response latencies at the IAT which does not disentangle different parameters (Greenwald et al., 2003). At the time, we argued that these results indicated that participants who actually performed the approach-avoidance training acquired a stronger evaluative response. Could it be that the difference we observed was the results of participants being more depleted in the repetition-based condition instead? We consider this interpretation extremely unlikely. If a repetition-based paradigm depleted the cognitive resources of the participants and if this depletion was the reason we found a difference in terms of D scores between our two approach-avoidance procedures, this would not explain why the differences between approach-avoidance training and approach-avoidance instructions persisted a day after the training. We consider extremely unlikely the hypothesis that an approach-avoidance training is a procedure so resource-consuming that its depleting effect would last up to a day after the training. Moreover, even if true, we should have been able to observe a more pronounced difference the first day than the second, as resources would gradually replenish.

Even if our results do not perfectly align with the ones described by Smith et al. (2019), the MPT methodology offers a more granular perspective on the expression of indirect evaluative responses. We think such methodology should help researchers better understand the different processes involve in the expression of an evaluative response. Indeed, compared to the factors relevant to the learning of evaluative responses, we have a lot to learn regarding the factors impacting the expression of indirect evaluative responses. With this idea in mind, the MPT methodology should be a useful tool (see Hütter & Klauer, 2016).

Methodological Limitations and Ways Forwards

As in any work, the present dissertation suffers from some limitations. The discussion provides the opportunity to return to them and to offer perspectives for future research. A first one has to do with the material we used in our experiments. A reasonable question to ask would be the one of the social nature of our stimuli. The logos that we used (Experiments 3 & 4) or the Luupite and Niffite lists of words (Experiments 1a, 1b, 5–7; Gregg et al., 2006) are far away from the initial focus of social psychology; these stimuli could hardly be considered as social stimuli. One could argue that the limited ecological validity of this line of research could pose problems in terms of ecological validity. We totally agree with this argument and consider that further study should progressively adopt a wider range of stimuli. In a reassuring way, however, it seems that the results of work focusing on the difference between approach-avoidance procedures with non-social and social stimuli align (e.g., the research on IAT; see Experiments 6 and 7 and Van Dessel et al., 2020). It seems that investigating the differences between repetition- and instructions-based with novel or known stimuli yield to the same result.

Another limitation of the present work is our focus on the repetition at a procedure level. Research on repetition distinguish at least two conceptually distinct ways to conceptualize repetition. The first one refers to the repetition of an event in a short period of time (as in many approach-avoidance trials), the second the repetition of an event over time (as in several approach-avoidance training distributed over time). In the present dissertation, we focused on the first type. The literature shows that the strength of acquired evaluative response can increase with the number of trials in a repetition task (Bar-Anan et al., 2010;

Woud et al., 2011).⁴⁵ It also shows that the repetition of the learning over time—also known as distributed learning—can increase the strength of acquired evaluative responses (Richter & Gast, 2017). Even if the Hebbian's rule described in associative processes maps onto these two conceptualizations, we focused on the repetition of an experience during a single experimental procedure. This was motivated by the emergence of instructions-based paradigms which raise the question of the difference between repetition- and instructions-based procedure. Yet, we think that comprehensively assessing the effect of repeated experience cannot be done without also investigating the effects of distributed learning.

On a more methodological note, some of our non-significant findings could have been more informative. In this dissertation, we adopted a NHST approach to test our hypotheses. This statistical approach is probably the most common in psychology, but it sometimes presents a problem: When a statistical result is not significant, it is hard to conclude on anything. Yet, alternatives making non-significant findings informative do exist (Lakens et al., 2018). Bayesian statistics (Morey & Rouder, 2018; Nalborczyk et al., 2019) or equivalence testing (Lakens, 2017) are some of them. In the present dissertation, we strictly adhered to our pre-registered analysis and did not use these procedures—unless a different analysis improved the consistency across experiments while supporting the same conclusion as the preregistered analysis. We think that the adoption of these alternative methods are promising to take full advantage of the empirical work conducted (e.g., Van Dessel et al., 2020).

⁴⁵ As discussed by Gast (2018), while repetition can increase the strength of indirect evaluative response, it sometimes does not. In the case of Evaluative Conditioning, Bar Anan et al. (2010) showed that the effect of the number of trials on evaluative responses decreased when participants had a perfect memory of the pairing in the procedure.

Finally, we restricted ourselves in the recruitment of either French, British, and American populations. While we tried to be as inclusive as we could regarding the screening criteria of our studies (in terms of gender, age, or economic status), our participants remain people from industrialized western countries (see Henrich et al., 2010a). Conducting research with the help of a collaborative network such as the Psychological Science Accelerator (Moshontz et al., 2018) or StudySwap (Chartier et al., 2018) could improve the external validity of our studies.

Concluding Thoughts

The work we undertook in this manuscript started as a very simple question. How much-repeated experience contributes to the acquisition of indirect evaluative responses? We think that this question is an important question to ask with regards to the recent advances of the implicit social cognition literature. In a nutshell, during the last decade, researchers showed that merely giving someone the instructions of procedures traditionally used to form indirect evaluative responses was. These results could be relatively surprising because one of the main ingredients of these procedures was the repetition of an event—an ingredient whose importance had been highlighted during the early days of implicit social cognition. Could the paradigms used in implicit social cognition be reduced to the effect of their instructions?

To tackle this question, we used a procedure where the repeated experience was movements of approach and avoidance. This type of paradigm—approach-avoidance training—has been used to tackle a wide range of problems, such as attitudes toward math, intergroup relationship, or addiction. We first used the latest advances in approach-avoidance movement implementation to develop an approach-avoidance training procedure and then compared it to an approach-avoidance instruction procedure. Across five different experiments, adopting different designs and investigating different evaluative responses in different contexts, we collected partial support for the idea of a contribution of repeated

experience: Some, but not all, experiments revealed a significant difference in terms of magnitude between the indirect evaluative responses acquired as a result of the two approach-avoidance procedures. Regarding the modern theories of implicit social cognition, these results were more easily explained by some theories than others. Specifically, our results were more consistent with propositional model of evaluations than with dual models. But beside the contribution to the theories of implicit social cognition, the research that we conducted allows us to make recommendations on the use of approach-avoidance procedures. Overall, despite the robustness of approach-avoidance instructions procedures when it comes to the learning of indirect evaluative response, approach-avoidance training appears to be a superior. It is important to acknowledge that this difference in favor of approach-avoidance training procedure seems to be a small effect. Our recommendations regarding the use of approach-avoidance training rather than an instructions-based procedure also relies on the naive beliefs people can have of these procedures. The mechanisms accounting for this superiority of approach-avoidance training remain to be investigated.

References

- Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, 84(5), 888-918.
<https://doi.org/10.1037/0033-2909.84.5.888>
- Allport, G. W. (1935). Attitudes. In *A Handbook of Social Psychology* (p. 798-844). Clark University Press.
- Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science*, 12(5), 385-390. <https://doi.org/10.1111/1467-9280.00371>
- Amodio, D. M., & Mendoza, S. A. (2010). Implicit intergroup bias: Cognitive, affective, and motivational underpinnings. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition* (p. 353-374). Guilford Press.
- Anvari, F., & Lakens, D. (2021). Using anchor-based methods to determine the smallest effect size of interest. *Journal of Experimental Social Psychology*, 96, 104159.
<https://doi.org/10.1016/j.jesp.2021.104159>
- Arkes, H. R., & Tetlock, P. E. (2004). Attributions of implicit prejudice, or “Would Jesse Jackson “fail” the Implicit Association Test?” *Psychological Inquiry*, 15(4), 257-278.
https://doi.org/10.1207/s15327965pli1504_01
- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63(4), 596-612. <https://doi.org/10.1037/0022-3514.63.4.596>

- Asendorpf, J. B., Banse, R., & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology*, 83(2), 380-393. <https://doi.org/10.1037/0022-3514.83.2.380>
- Aubé, B., Rougier, M., Muller, D., Ric, F., & Yzerbyt, V. (2019). The online-VAAST: A short and online tool to measure spontaneous approach and avoidance tendencies. *Acta Psychologica*, 201, 102942. <https://doi.org/10.1016/j.actpsy.2019.102942>
- Baddeley, A. D. (2010). Working memory. *Current Biology*, 20(4), R136-R140. <https://doi.org/10.1016/j.cub.2009.12.014>
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of Learning and Motivation* (Vol. 8, p. 47-89). Elsevier. [https://doi.org/10.1016/S0079-7421\(08\)60452-1](https://doi.org/10.1016/S0079-7421(08)60452-1)
- Bar-Anan, Y., De Houwer, J., & Nosek, B. A. (2010). Evaluative conditioning and conscious knowledge of contingencies: A correlational investigation with large samples. *Quarterly Journal of Experimental Psychology*, 63(12), 2313-2335. <https://doi.org/10.1080/17470211003802442>
- Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (p. 1-40). Erlbaum.
- Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (p. 361-382). The Guilford Press.
- Bargh, J. A. (2017). *Before you know it: The unconscious reasons we do what we do*. Touchstone.
- Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2010). A sketch of the Implicit Relational Assessment Procedure (IRAP) and the Relational Elaboration and

- Coherence (REC) Model. *The Psychological Record*, 60(3), 527-542.
<https://doi.org/10.1007/BF03395726>
- Batailler, C., Muller, D., Nurra, C., Rougier, M., & Trouilloud, D. (2021). Math approach training changes implicit identification with math: A close preregistered replication. *Journal of Experimental Social Psychology*, 92, 104059.
<https://doi.org/10.1016/j.jesp.2020.104059>
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*. <http://arxiv.org/abs/1506.04967>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using **lme4**. *Journal of Statistical Software*, 67(1).
<https://doi.org/10.18637/jss.v067.i01>
- Becker, E. S., Barth, A., Smits, J. A. J., Beisel, S., Lindenmeyer, J., & Rinck, M. (2019). Positivity-approach training for depressive symptoms: A randomized controlled trial. *Journal of Affective Disorders*, 245, 297-304.
<https://doi.org/10.1016/j.jad.2018.11.042>
- Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., Bollen, K. A., Brembs, B., Brown, L., Camerer, C., Cesarini, D., Chambers, C. D., Clyde, M., Cook, T. D., De Boeck, P., Dienes, Z., Dreber, A., Easwaran, K., Efferson, C., ... Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour*, 2(1), 6-10. <https://doi.org/10.1038/s41562-017-0189-z>
- Berke, R. L. (2000, September 12). The 2000 campaign: The ad campaign; Democrats see, and smell, rats in g.o.p. Ad. *The New York Times*.
<https://www.nytimes.com/2000/09/12/us/the-2000-campaign-the-ad-campaign-democrats-see-and-smell-rats-in-gop-ad.html>

- Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, 6(3), 242-261.
https://doi.org/10.1207/S15327957PSPR0603_8
- Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *American Psychologist*, 61(1), 27.
- Boffo, M., Willemen, R., Pronk, T., Wiers, R. W., & Dom, G. (2017). Effectiveness of two web-based cognitive bias modification interventions targeting approach and attentional bias in gambling problems: Study protocol for a pilot randomised controlled trial. *Trials*, 18(1). <https://doi.org/10.1186/s13063-017-2190-2>
- Boffo, M., Zerhouni, O., Gronau, Q. F., van Beek, R. J. J., Nikolaou, K., Marsman, M., & Wiers, R. W. (2019). Cognitive bias modification for behavior change in alcohol and smoking addiction: Bayesian meta-analysis of individual participant data. *Neuropsychology Review*. <https://doi.org/10.1007/s11065-018-9386-4>
- Bohner, G., & Dickel, N. (2011). Attitudes and attitude change. *Annual Review of Psychology*, 62(1), 391-417. <https://doi.org/10.1146/annurev.psych.121208.131609>
- Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, 106(2), 265-289. <https://doi.org/10.1037/0033-2909.106.2.265>
- Braisby, N., Franks, B., & Hampton, J. (1996). Essentialism, word use, and concepts. *Cognition*, 59(3), 247-274. [https://doi.org/10.1016/0010-0277\(95\)00698-2](https://doi.org/10.1016/0010-0277(95)00698-2)
- Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., Grange, J. A., Perugini, M., Spies, J. R., & van 't Veer, A. (2014). The replication recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*, 50, 217-224. <https://doi.org/10.1016/j.jesp.2013.10.005>

- Braud, M. (2020). *Failure to find a superior effect of approach-avoidance training over mere instructions in an impression formation paradigm* [Master dissertation]. Université Grenoble Alpes.
- Brehm, J. W. (1956). Postdecision changes in the desirability of alternatives. *The Journal of Abnormal and Social Psychology*, 52(3), 384-389. <https://doi.org/10.1037/h0041006>
- Brewer, M. B., & Kramer, R. M. (1985). The psychology of intergroup attitudes and behavior. *Annual Review of Psychology*, 36(1), 219-243. <https://doi.org/10.1146/annurev.ps.36.020185.001251>
- Cacioppo, J. T., Priester, J. R., & Berntson, G. G. (1993). Rudimentary determinants of attitudes: II. Arm flexion and extension have differential effects on attitudes. *Journal of Personality and Social Psychology*, 65(1), 5-17. <https://doi.org/10.1037/0022-3514.65.1.5>
- Calanchini, J., & Sherman, J. W. (2013). Implicit attitudes reflect associative, non-associative, and non-attitudinal processes: Implicit attitudes reflect non-attitudinal processes. *Social and Personality Psychology Compass*, 7(9), 654-667. <https://doi.org/10.1111/spc3.12053>
- Carraro, L., Gawronski, B., & Castelli, L. (2010). Losing on all fronts: The effects of negative versus positive person-based campaigns on implicit and explicit evaluations of political candidates. *British Journal of Social Psychology*, 49(3), 453-470. <https://doi.org/10.1348/014466609X468042>
- Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). correcting for bias in psychology: A comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological Science*, 2(2), 115-144. <https://doi.org/10.1177/2515245919847196>

Centerbar, D. B., & Clore, G. L. (2006). Do approach-avoidance actions create attitudes?

Psychological science, 17(1), 22-29.

Chartier, C. R., Riegelman, A., & McCarthy, R. J. (2018). StudySwap: A platform for

interlab replication, collaboration, and resource exchange. *Advances in Methods and Practices in Psychological Science*, 1(4), 574-579.

<https://doi.org/10.1177/2515245918808767>

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate

behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, 25(2), 215-224. <https://doi.org/10.1177/0146167299025002007>

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. ed.). Routledge.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.

<https://doi.org/10.1037/0033-2909.112.1.155>

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005).

Separating multiple processes in implicit social cognition: The Quad Model of Implicit Task performance. *Journal of Personality and Social Psychology*, 89(4), 469-487. <https://doi.org/10.1037/0022-3514.89.4.469>

Corneille, O., & Hütter, M. (2020). Implicit? What do you mean? A Comprehensive Review

of the Delusive Implicitness Construct in Attitude Research. *Personality and Social Psychology Review*, 108886832091132. <https://doi.org/10.1177/1088868320911325>

Corneille, O., & Mertens, G. (2020). Behavioral and physiological evidence challenges the

automatic acquisition of Evaluations. *Current Directions in Psychological Science*, 29(6), 569-574. <https://doi.org/10.1177/0963721420964111>

Corneille, O., Mierop, A., Stahl, C., & Hütter, M. (2019). Evidence suggestive of

uncontrollable attitude acquisition replicates in an instructions-based evaluative conditioning paradigm: Implications for associative attitude acquisition. *Journal of*

- Experimental Social Psychology*, 85, 103841.
<https://doi.org/10.1016/j.jesp.2019.103841>
- Corneille, O., & Stahl, C. (2019). Associative attitude learning: A closer look at evidence and how it relates to attitude models. *Personality and Social Psychology Review*, 23(2), 161-189. <https://doi.org/10.1177/1088868318763261>
- Correll, J., Park, B., Judd, C. M., Wittenbrink, B., Sadler, M. S., & Keesee, T. (2007). Across the thin blue line: Police officers and racial bias in the decision to shoot. *Journal of Personality and Social Psychology*, 92(6), 1006-1023. <https://doi.org/10.1037/0022-3514.92.6.1006>
- Courset, R., Rougier, M., Palluel-Germain, R., Smeding, A., Jonte, J. M., Chauvin, A., & Muller, D. (2018). The Caucasian and North African French Faces (CaNAFF): A Face Database. *International Review of Social Psychology*, 31(1).
<https://doi.org/10.5334/irsp.179>
- Cumming, G. (2008). Replication and p intervals: P values predict the future only vaguely, but confidence intervals do much Better. *Perspectives on Psychological Science*, 3(4), 286-300. <https://doi.org/10.1111/j.1745-6924.2008.00079.x>
- Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science*, 12(2), 163-170. <https://doi.org/10.1111/1467-9280.00328>
- Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, 81(5), 800-814.
<https://doi.org/10.1037/0022-3514.81.5.800>
- De Houwer, J. (2003). The Extrinsic Affective Simon Task. *Experimental Psychology*, 50(2), 77-85. <https://doi.org/10.1026//1618-3169.50.2.77>

- De Houwer, J. (2006a). What are implicit measures and why are we using them? In R. W. Wiers & A. W. Stacy (Eds.), *The handbook of implicit cognition and addiction* (p. 11-28). Sage.
- De Houwer, J. (2006b). Using the Implicit Association Test does not rule out an impact of conscious propositional knowledge on evaluative conditioning. *Learning and Motivation*, 37(2), 176-187. <https://doi.org/10.1016/j.lmot.2005.12.002>
- De Houwer, J. (2007). A conceptual and theoretical analysis of evaluative conditioning. *The Spanish Journal of Psychology*, 10(2), 230-241. <https://doi.org/10.1017/S1138741600006491>
- De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, 37(1), 1-20. <https://doi.org/10.3758/LB.37.1.1>
- De Houwer, J. (2014). A propositional model of implicit evaluation. *Social and Personality Psychology Compass*, 8(7), 342-353. <https://doi.org/10.1111/spc3.12111>
- De Houwer, J. (2018). Propositional models of evaluative conditioning. *Social Psychological Bulletin*, 13(3). <https://doi.org/10.5964/spb.v13i3.28046>
- De Houwer, J. (2019). Implicit bias is behavior: a functional-cognitive perspective on implicit bias. *Perspectives on Psychological Science*, 174569161985563. <https://doi.org/10.1177/1745691619855638>
- De Houwer, J., Gawronski, B., & Barnes-Holmes, D. (2013). A functional-cognitive framework for attitude research. *European Review of Social Psychology*, 24(1), 252-287. <https://doi.org/10.1080/10463283.2014.892320>
- De Houwer, J., Heider, N., Spruyt, A., Roets, A., & Hughes, S. (2015). The relational responding task: Toward a new implicit measure of beliefs. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00319>

- De Houwer, J., & Hughes, S. (2016). Evaluative conditioning as a symbolic phenomenon: on the relation between evaluative conditioning, evaluative conditioning via instructions, and persuasion. *Social Cognition*, 34(5), 480-494.
<https://doi.org/10.1521/soco.2016.34.5.480>
- De Houwer, J., Thomas, S., & Baeyens, F. (2001). Association learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin*, 127(6), 853-869. <https://doi.org/10.1037/0033-2909.127.6.853>
- de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, 47(1), 1-12.
<https://doi.org/10.3758/s13428-014-0458-y>
- de Leeuw, J. R., & Motz, B. A. (2016). Psychophysics in a web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behavior Research Methods*, 48(1), 1-12. <https://doi.org/10.3758/s13428-015-0567-2>
- DeMaio, T. J. (1984). Social desirability and survey measurement: A review. In C. F. Turner & E. Martin (Eds.), *Surveying Subjective Phenomena, Volume 2* (p. 257-282). Russell Sage Foundation. <http://www.jstor.org/stable/10.7758/9781610447003.16>
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56(1), 5-18.
<https://doi.org/10.1037/0022-3514.56.1.5>
- Dotsch, R., & Todorov, A. (2012). Reverse correlating social face perception. *Social Psychological and Personality Science*, 3(5), 562-571.
<https://doi.org/10.1177/1948550611430272>

- Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, 82(1), 62-68.
<https://doi.org/10.1037/0022-3514.82.1.62>
- Duckworth, K. L., Bargh, J. A., Garcia, M., & Chaiken, S. (2002). The automatic evaluation of novel stimuli. *Psychological Science*, 13(6), 513-519. <https://doi.org/10.1111/1467-9280.00490>
- Eager, C., & Roy, J. (2017). mixed effects models are sometimes terrible. *arXiv preprint arXiv:1701.04858*. <https://arxiv.org/abs/1701.04858>
- Eder, A. B., Krishna, A., & Van Dessel, P. (2019). Operant evaluative conditioning. *Journal of Experimental Psychology: Animal Learning and Cognition*, 45(1), 102-110.
<https://doi.org/10.1037/xan0000189>
- Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In *Advances in experimental social psychology* (Vol. 23, p. 75-109).
- 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>

- Fazio, R. H., & Towles-Schwen, T. (1999). The MODE model of attitude-behavior processes. In S. Chaiken & Y. Trope (Eds.), *Dual-Processes Theories in Social Psychology*. Guilford Press.
- Ferentzi, H., Scheibner, H., Wiers, R. W., Becker, E. S., Lindenmeyer, J., Beisel, S., & Rinck, M. (2018). Retraining of automatic action tendencies in individuals with obesity: A randomized controlled trial. *Appetite*, 126, 66-72.
<https://doi.org/10.1016/j.appet.2018.03.016>
- Ferguson, C. J., & Heene, M. (2012). A Vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science*, 7(6), 555-561. <https://doi.org/10.1177/1745691612459059>
- Ferguson, M. J. (2008). On becoming ready to pursue a goal you don't know you have: Effects of nonconscious goals on evaluative readiness. *Journal of Personality and Social Psychology*, 95(6), 1268-1294. <https://doi.org/10.1037/a0013263>
- Festinger, L. (1968). *A theory of cognitive dissonance*. Row, Peterson.
- Fidler, F. (2002). The fifth edition of the APA publication manual: Why its statistics recommendations are so controversial. *Educational and Psychological Measurement*, 62(5), 749-770. <https://doi.org/10.1177/001316402236876>
- Fisk, A. D., & Schneider, W. (1984). Memory as a function of attention, level of processing, and automatization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(2), 181-197. <https://doi.org/10.1037/0278-7393.10.2.181>
- 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>

- Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B. A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology*. <https://doi.org/10.1037/pspa0000160>
- Fort, K., Adda, G., & Cohen, K. B. (2011). Amazon Mechanical Turk: Gold mine or coal mine? *Computational Linguistics*, 37(2), 413-420.
https://doi.org/10.1162/COLI_a_00057
- Funder, D. C., Levine, J. M., Mackie, D. M., Morf, C. C., Sansone, C., Vazire, S., & West, S. G. (2014). Improving the dependability of research in personality and social psychology: Recommendations for Research and Educational Practice. *Personality and Social Psychology Review*, 18(1), 3-12.
<https://doi.org/10.1177/1088868313507536>
- Gast, A. (2018). A declarative memory model of evaluative conditioning. *Social Psychological Bulletin*, 13(3). <https://doi.org/10.5964/spb.v13i3.28590>
- Gast, A., & De Houwer, J. (2012). Evaluative conditioning without directly experienced pairings of the conditioned and the unconditioned stimuli. *Quarterly Journal of Experimental Psychology*, 65(9), 1657-1674.
<https://doi.org/10.1080/17470218.2012.665061>
- Gawronski, B., & 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. (2007). Unraveling the processes underlying evaluation: Attitudes from the perspective of the APE model. *Social Cognition*, 25(5), 687-717. <https://doi.org/10.1521/soco.2007.25.5.687>
- 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. (2014a). The associative-propositional evaluation model: operating principles and operating conditions of evaluation. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (p. 188-203). The Guilford Press.
- Gawronski, B., & Bodenhausen, G. V. (2014b). Implicit and explicit evaluation: a brief review of the associative-propositional evaluation model: APE Model. *Social and Personality Psychology Compass*, 8(8), 448-462. <https://doi.org/10.1111/spc3.12124>
- Gawronski, B., & Bodenhausen, G. V. (2015). Theory evaluation. In B. Gawronski & G. V. Bodenhausen (Eds.), *Theory and Explanation in Social Psychology* (p. 3-23). Guilford Publications.
- Gawronski, B., & Bodenhausen, G. V. (2018). Evaluative conditioning from the perspective of the associative-propositional evaluation model. *Social Psychological Bulletin*, 13(3). <https://doi.org/10.5964/spb.v13i3.28024>
- Gawronski, B., De Houwer, J., & Sherman, J. W. (2020). Twenty-five years of research using implicit measures. *Social Cognition*.
- Gawronski, B., & Hahn, A. (2019). Implicit measures: Procedures, use, and interpretation. In H. Blanton, J. M. LaCroix, & G. D. Webster (Eds.), *Measurement in social psychology* (p. 29-55). Taylor & Francis.
- Gawronski, B., & Payne, B. K. (Eds.). (2010). *Handbook of implicit social cognition: Measurement, theory, and applications*. Guilford Press.
- Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, 102, 74-78.
<https://doi.org/10.1016/j.paid.2016.06.069>

Giner-Sorolla, R. (2016). Approaching a fair deal for significance and other concerns.

Journal of Experimental Social Psychology, 65, 1-6.

<https://doi.org/10.1016/j.jesp.2016.01.010>

Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2014). Measures of implicit attitudes may conceal differences in implicit associations: The case of antiaging bias. *Social Psychological and Personality Science*, 5(3), 271-278.

<https://doi.org/10.1177/1948550613499239>

Gonzales, P., Guzman, J. C., Partelow, L., Pahlke, E., Jocelyn, L., Kastberg, D., & Williams, T. (2004). *Highlights from the Trends in International Mathematics and Science Study (TIMSS) 2003* [Data set]. American Psychological Association.

<https://doi.org/10.1037/e672832007-002>

Gorn, G. J. (1982). The effects of music in advertising on choice behavior: a classical conditioning approach. *Journal of Marketing*, 46(1), 94-101.

<https://doi.org/10.1177/002224298204600109>

Greenberg, R. P., Constantino, M. J., & Bruce, N. (2006). Are patient expectations still relevant for psychotherapy process and outcome? *Clinical Psychology Review*, 26(6), 657-678. <https://doi.org/10.1016/j.cpr.2005.03.002>

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1), 4-27. <https://doi.org/10.1037/0033-295X.102.1.4>

Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2015). Statistically small effects of the Implicit Association Test can have societally large effects. *Journal of Personality and Social Psychology*, 108(4), 553-561. <https://doi.org/10.1037/pspa0000016>

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., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464-1480. <https://doi.org/10.1037/0022-3514.74.6.1464>
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85(2), 197-216. <https://doi.org/10.1037/0022-3514.85.2.197>
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, 97(1), 17-41. <https://doi.org/10.1037/a0015575>
- Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: Asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology*, 90(1), 1-20. <https://doi.org/10.1037/0022-3514.90.1.1>
- Gregory, R. L., & Zangwill, O. L. (Eds.). (1987). *The Oxford companion to the mind*. Oxford University Press.
- Hahn, A., & Goedderz, A. (2020). Trait-unconsciousness, state-unconsciousness, preconsciousness, and social miscalibration in the context of implicit evaluation. *Social Cognition*, 38(Supplement), s115-s134. <https://doi.org/10.1521/soco.2020.38.sup.s115>
- Hahn, A. M., Simons, R. M., Simons, J. S., Wiers, R. W., & Welker, L. E. (2019). Can cognitive bias modification simultaneously target two behaviors? Approach bias

- retraining for alcohol and condom use. *Clinical Psychological Science*, 216770261983457. <https://doi.org/10.1177/2167702619834570>
- Hebb, D. O. (1949). *The organization of behavior*. Wiley & Sons.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010a). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2-3), 61-83.
<https://doi.org/10.1017/S0140525X0999152X>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010b). Most people are not WEIRD. *Nature*, 466(7302), 29-29. <https://doi.org/10.1038/466029a>
- Herring, D. R., White, K. R., Jabeen, L. N., Hinojos, M., Terrazas, G., Reyes, S. M., Taylor, J. H., & Crites, S. L. (2013). On the automatic activation of attitudes: A quarter century of evaluative priming research. *Psychological Bulletin*, 139(5), 1062-1089.
<https://doi.org/10.1037/a0031309>
- Heycke, T., & Gawronski, B. (2020). Co-occurrence and relational information in evaluative learning: A multinomial modeling approach. *Journal of Experimental Psychology: General*, 149(1), 104-124. <https://doi.org/10.1037/xge0000620>
- Heycke, T., Gehrman, 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>
- Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, 16(2), 96-101.
<https://doi.org/10.3758/BF03202365>
- Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin*, 136(3), 390-421.
<https://doi.org/10.1037/a0018916>

- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality and Social Psychology Bulletin*, 31(10), 1369-1385. <https://doi.org/10.1177/0146167205275613>
- 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>
- Hu, X., Gawronski, B., & Balas, R. (2017). Propositional versus dual-process accounts of evaluative conditioning: II. The effectiveness of counter-conditioning and counter-instructions in changing implicit and explicit evaluations. *Social Psychological and Personality Science*, 8(8), 858-866. <https://doi.org/10.1177/1948550617691094>
- Hughes, S., Barnes-Holmes, D., & De Houwer, J. (2011). The dominance of associative theorizing in implicit attitude research: Propositional and behavioral alternatives. *The Psychological Record*, 61(3), 465-496. <https://doi.org/10.1007/BF03395772>
- Hummel, J. E. (2010). Symbolic versus associative learning. *Cognitive Science*, 34(6), 958-965. <https://doi.org/10.1111/j.1551-6709.2010.01096.x>
- 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. (2018). Dissociating controllable and uncontrollable effects of affective stimuli on attitudes and consumption. *Journal of Consumer Research*, 45(2), 320-349. <https://doi.org/10.1093/jcr/ucx124>

- Imhoff, R., & Banse, R. (2009). Ongoing victim suffering increases prejudice: The case of secondary anti-semitism. *Psychological Science*, 20(12), 1443-1447.
<https://doi.org/10.1111/j.1467-9280.2009.02457.x>
- Ioannidis, J. P. A. (2008). Why most discovered true associations are inflated: *Epidemiology*, 19(5), 640-648. <https://doi.org/10.1097/EDE.0b013e31818131e7>
- Jacobus, J., Taylor, C. T., Gray, K. M., Meredith, L. R., Porter, A. M., Li, I., Castro, N., & Squeglia, L. M. (2018). A multi-site proof-of-concept investigation of computerized approach-avoidance training in adolescent cannabis users. *Drug and Alcohol Dependence*, 187, 195-204. <https://doi.org/10.1016/j.drugalcdep.2018.03.007>
- Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, 110(3), 306-340. <https://doi.org/10.1037/0096-3445.110.3.306>
- Jeffreys, H. (1961). *Theory of probability*. Oxford University Press.
- Jones, C. R., Vilensky, M. R., Vasey, M. W., & Fazio, R. H. (2013). Approach behavior can mitigate predominately univalent negative attitudes: Evidence regarding insects and spiders. *Emotion*, 13(5), 989-996. <https://doi.org/10.1037/a0033164>
- Jones, E. E., & Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin*, 76(5), 349-364.
<https://doi.org/10.1037/h0031617>
- Judd, C. M., McClelland, G. H., & Ryan, C. S. (2017). *Data analysis: A model comparison approach to regression, ANOVA, and beyond* (Third Edition). Routledge, Taylor & Francis Group.
- Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely

- ignored problem. *Journal of Personality and Social Psychology*, 103(1), 54-69.
<https://doi.org/10.1037/a0028347>
- Kahneman, D. (2011). *Thinking, fast and slow* (1st ed). Farrar, Straus and Giroux.
- Kakoschke, N., Kemps, E., & Tiggemann, M. (2017). Approach bias modification training and consumption: A review of the literature. *Addictive Behaviors*, 64, 21-28.
<https://doi.org/10.1016/j.addbeh.2016.08.007>
- Kawakami, K., Phillips, C. E., Steele, J. R., & Dovidio, J. F. (2007). (Close) distance makes the heart grow fonder: Improving implicit racial attitudes and interracial interactions through approach behaviors. *Journal of Personality and Social Psychology*, 92(6), 957-971. <https://doi.org/10.1037/0022-3514.92.6.957>
- Kawakami, K., Steele, J. R., Cifa, C., Phillips, C. E., & Dovidio, J. F. (2008). Approaching math increases math=me and math=pleasant. *Journal of Experimental Social Psychology*, 44(3), 818-825. <https://doi.org/10.1016/j.jesp.2007.07.009>
- Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, 28(2), 107-128. <https://doi.org/10.1037/h0034225>
- Krieglmeyer, R., & Deutsch, R. (2010). Comparing measures of approach–avoidance behaviour: The manikin task vs. two versions of the joystick task. *Cognition & Emotion*, 24(5), 810-828. <https://doi.org/10.1080/02699930903047298>
- Krieglmeyer, R., & Sherman, J. W. (2012). Disentangling stereotype activation and stereotype application in the stereotype misperception task. *Journal of Personality and Social Psychology*, 103(2), 205-224. <https://doi.org/10.1037/a0028764>
- Krosnick, J. A., Judd, C. M., & Wittenbrink, B. (2005). The measurement of attitudes. In D. Albarracín & B. T. Johnson (Eds.), *The handbook of attitudes*. Psychology Press.

- 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>
- Kurdi, B., Seitchik, A. E., Axt, J. R., Carroll, T. J., Karapetyan, A., Kaushik, N., Tomezsko, D., Greenwald, A. G., & Banaji, M. R. (2018). Relationship between the Implicit Association Test and intergroup behavior: A meta-analysis. *American Psychologist*. <https://doi.org/10.1037/amp0000364>
- Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J.-E. L., Joy-Gaba, J. A., Ho, A. K., Teachman, B. A., Wojcik, S. P., Koleva, S. P., Frazier, R. S., Heiphetz, L., Chen, E. E., Turner, R. N., Haidt, J., Kesebir, S., Hawkins, C. B., Schaefer, H. S., Rubichi, S., ... Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, 143(4), 1765-1785. <https://doi.org/10.1037/a0036260>
- Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., Calanchini, J., Xiao, Y. J., Pedram, C., Marshburn, C. K., Simon, S., Blanchar, J. C., Joy-Gaba, J. A., Conway, J., Redford, L., Klein, R. A., Roussos, G., Schellhaas, F. M. H., Burns, M., ... Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, 145(8), 1001-1016. <https://doi.org/10.1037/xge0000179>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00863>
- Lakens, D. (2017). Equivalence tests: A Practical Primer for *t* Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, 8(4), 355-362. <https://doi.org/10.1177/1948550617697177>

- Lakens, D., & Evers, E. R. K. (2014). Sailing from the seas of chaos into the corridor of stability: Practical recommendations to increase the informational value of studies. *Perspectives on Psychological Science*, 9(3), 278-292.
<https://doi.org/10.1177/1745691614528520>
- Lakens, D., McLatchie, N., Isager, P. M., Scheel, A. M., & Dienes, Z. (2018). Improving inferences about null effects with bayes factors and equivalence tests. *The Journals of Gerontology: Series B*. <https://doi.org/10.1093/geronb/gby065>
- Lang, P., & Bradley, M. M. (2007). The International Affective Picture System (IAPS) in the study of emotion and attention. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of emotion elicitation and assessment* (p. 29-46). Oxford University Press.
- LaPiere, R. T. (1934). Attitudes vs. Actions. *Social Forces*, 13(2), 230-237.
<https://doi.org/10.2307/2570339>
- 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](https://doi.org/10.1016/0005-7967(75)90026-1)
- Lovakov, A., & Agadullina, E. R. (2021). Empirically derived guidelines for effect size interpretation in social psychology. *European Journal of Social Psychology*, ejsp.2752. <https://doi.org/10.1002/ejsp.2752>
- Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology*, 81(5), 842-855.
<https://doi.org/10.1037/0022-3514.81.5.842>
- Magnusson, K. (2021). *Interpreting cohen's d effect size: An interactive visualization* (2.5.1) [Computer software]. <https://rpsychologist.com/cohend/>

- Mayo, D. G., & Spanos, A. (2006). Severe testing as a basic concept in a Neyman–Pearson philosophy of induction. *The British Journal for the Philosophy of Science*, 57(2), 323-357. <https://doi.org/10.1093/bjps/axl003>
- McConahay, J. B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (p. 91-125). Academic Press.
- McConnell, A. R., & Rydell, R. J. (2014). The systems of evaluation model: A dual-systems approach to attitudes. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-Process Theories of the Social Mind* (p. 204-218). The Guilford Press.
- McConnell, A. R., Rydell, R. J., Strain, L. M., & Mackie, D. M. (2008). Forming implicit and explicit attitudes toward individuals: Social group association cues. *Journal of Personality and Social Psychology*, 94(5), 792-807. <https://doi.org/10.1037/0022-3514.94.5.792>
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81-97. <https://doi.org/10.1037/h0043158>
- Miller, G. A. (2003). The cognitive revolution: A historical perspective. *Trends in Cognitive Sciences*, 7(3), 141-144. [https://doi.org/10.1016/S1364-6613\(03\)00029-9](https://doi.org/10.1016/S1364-6613(03)00029-9)
- 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>
- 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. (2018). *BayesFactor: Computation of bayes factors for common designs* [Manual]. <https://CRAN.R-project.org/package=BayesFactor>

- Moshontz, H., Campbell, L., Ebersole, C. R., IJzerman, H., Urry, H. L., Forscher, P. S., Grahe, J. E., McCarthy, R. J., Musser, E. D., Antfolk, J., Castille, C. M., Evans, T. R., Fiedler, S., Flake, J. K., Forero, D. A., Janssen, S. M. J., Keene, J. R., Protzko, J., Aczel, B., ... Chartier, C. R. (2018). The Psychological Science Accelerator: Advancing psychology through a distributed collaborative network. *Advances in Methods and Practices in Psychological Science*, 1(4), 501-515.
<https://doi.org/10.1177/2515245918797607>
- Nalborczyk, L., Batailler, C., Løevenbruck, H., Vilain, A., & Bürkner, P.-C. (2019). An Introduction to Bayesian Multilevel Models Using brms: A Case Study of Gender Effects on Vowel Variability in Standard Indonesian. *Journal of Speech, Language, and Hearing Research*, 62(5), 1225-1242. https://doi.org/10.1044/2018_JSLHR-S-18-0006
- Necka, E. A., Sokolowski, H. M., & Lyons, I. M. (2015). The role of self-math overlap in understanding math anxiety and the relation between math anxiety and performance. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01543>
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231-259.
<https://doi.org/10.1037//0033-295X.84.3.231>
- Nosek, B. A., & Banaji, M. R. (2001). The Go/No-Go Association Task. *Social Cognition*, 19(6), 625-666. <https://doi.org/10.1521/soco.19.6.625.20886>
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math \neq me. *Journal of Personality and Social Psychology*, 83(1), 44-59.
<https://doi.org/10.1037//0022-3514.83.1.44>
- Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., Bar-Anan, Y., Bergh, R., Cai, H., Gonsalkorale, K., Kesebir, S., Maliszewski, N., Neto, F., Olli, E.,

- Park, J., Schnabel, K., Shiomura, K., Tulbure, B. T., Wiers, R. W., ... Greenwald, A. G. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, 106(26), 10593-10597. <https://doi.org/10.1073/pnas.0809921106>
- OECD. (2015). *The ABC of Gender Equality in Education*. OECD Publishing. <https://doi.org/10.1787/9789264229945-en>
- Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17(11), 776-783. <https://doi.org/10.1037/h0043424>
- Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, 105(2), 171-192. <https://doi.org/10.1037/a0032734>
- Paladino, M.-P., & Castelli, L. (2008). On the immediate consequences of intergroup categorization: activation of approach and avoidance motor behavior toward ingroup and outgroup members. *Personality and Social Psychology Bulletin*, 34(6), 755-768. <https://doi.org/10.1177/0146167208315155>
- Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17, 22-27. <https://doi.org/10.1016/j.jbef.2017.12.004>
- Papies, E. K., Stroebe, W., & Aarts, H. (2009). Who likes it more? Restrained eaters' implicit attitudes towards food. *Appetite*, 53(3), 279-287. <https://doi.org/10.1016/j.appet.2009.07.001>

- 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>
- Payne, B. K., & Gawronski, B. (2010). A history of implicit social cognition: Where is it coming from? Where is it now? Where is it going. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (Vol. 1, p. 1-15). Guilford Press.
- Payne, B. K., Govorun, O., & Arbuckle, N. L. (2008). Automatic attitudes and alcohol: Does implicit liking predict drinking? *Cognition & Emotion*, 22(2), 238-271. <https://doi.org/10.1080/02699930701357394>
- Payne, B. K., & Lundberg, K. (2014). The Affect Misattribution Procedure: ten years of evidence on reliability, validity, and mechanisms. *Social and Personality Psychology Compass*, 8(12), 672-686. <https://doi.org/10.1111/spc3.12148>
- Perugini, M., Gallucci, M., & Costantini, G. (2018). A practical primer to power analysis for simple experimental designs. *International Review of Social Psychology*, 31(1). <https://doi.org/10.5334/irsp.181>
- Perugini, M., Richetin, J., & Zogmaister, C. (2010). Prediction of behavior. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (p. 225-277). Guilford Press.
- 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>
- Petty, R. E., & Briñol, P. (2006). A metacognitive approach to « implicit » and « explicit » evaluations: Comment on Gawronski and Bodenhausen (2006). *Psychological Bulletin*, 132(5), 740-744. <https://doi.org/10.1037/0033-2909.132.5.740>

- Petty, R. E., Briñol, P., & DeMarree, K. G. (2007). The Meta-Cognitive Model (MCM) of attitudes: Implications for attitude measurement, change, and strength. *Social Cognition*, 25(5), 657-686. <https://doi.org/10.1521/soco.2007.25.5.657>
- Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of persuasion. In R. E. Petty & J. T. Cacioppo, *Communication and Persuasion* (p. 1-24). Springer New York. https://doi.org/10.1007/978-1-4612-4964-1_1
- Phaf, R. H., Mohr, S. E., Rotteveel, M., & Wicherts, J. M. (2014). Approach, avoidance, and affect: A meta-analysis of approach-avoidance tendencies in manual reaction time tasks. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.00378>
- Phills, C. E., Kawakami, K., Tabi, E., Nadolny, D., & Inzlicht, M. (2011). Mind the gap: Increasing associations between the self and blacks with approach behaviors. *Journal of Personality and Social Psychology*, 100(2), 197-210. <https://doi.org/10.1037/a0022159>
- Popa-Roch, M.-A., & Delmas, F. (2011). Evidence for the salience asymmetries producing iat effects: replication and extension with a French-North African IAT. *Revue internationale de psychologie sociale*, 24(1), 37-61. Cairn.info.
- Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review*, 83(2), 157-171. <https://doi.org/10.1037/0033-295X.83.2.157>
- Rabbitt, P. (1979). How old and young subjects monitor and control responses for accuracy and speed. *British Journal of Psychology*, 70(2), 305-311. <https://doi.org/10.1111/j.2044-8295.1979.tb01687.x>
- Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, 7(4), 331-363. <https://doi.org/10.1037/1089-2680.7.4.331>

- Richetin, J., Costantini, G., Perugini, M., & Schönbrodt, F. (2015). Should we stop looking for a better scoring algorithm for handling implicit association test data? Test of the role of errors, extreme latencies treatment, scoring formula, and practice trials on reliability and validity. *PLOS ONE*, *10*(6), e0129601.
<https://doi.org/10.1371/journal.pone.0129601>
- 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>
- Rivers, A. M., & Hahn, A. (2019). What cognitive mechanisms do people reflect on when they predict IAT scores? *Personality and Social Psychology Bulletin*, *45*(6), 878-892.
<https://doi.org/10.1177/0146167218799307>
- Rohrer, J. M. (2018). Thinking clearly about correlations and causation: graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science*, 27-42. <https://doi.org/10.1177/2515245917745629>
- Roland, N., Mierop, A., Frenay, M., & Corneille, O. (2018). Field-Identification IAT predicts students' academic persistence over and above Theory of Planned Behavior constructs. *Frontline Learning Research*, *5*(1), 19-30.
<https://doi.org/10.14786/flr.v6i1.327>
- Rosenberg, M. J. (2009). The conditions and consequences of evaluation apprehension. In R. Rosenthal, R. L. Rosnow, & Alan E. Kazdin (Eds.), *Artifacts in Behavioral Research* (p. 211-263). Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780195385540.001.0001>
- Rothermund, K., & Wentura, D. (2004). Underlying processes in the implicit association test: dissociating salience from associations. *Journal of Experimental Psychology: General*, *133*(2), 139-165. <https://doi.org/10.1037/0096-3445.133.2.139>

- Rougier, M., Muller, D., Courset, R., Smeding, A., Devos, T., & Batailler, C. (2020). Toward the use of approach/avoidance tendencies as attitude measures: Individual- and group-level variability of the ingroup bias. *European Journal of Social Psychology*, ejsp.2653. <https://doi.org/10.1002/ejsp.2653>
- Rougier, M., Muller, D., Ric, F., Alexopoulos, T., Batailler, C., Smeding, A., & Aubé, B. (2018). A new look at sensorimotor aspects in approach/avoidance tendencies: The role of visual whole-body movement information. *Journal of Experimental Social Psychology*, 76, 42-53. <https://doi.org/10.1016/j.jesp.2017.12.004>
- Rougier, M., Schmitz, M., & Yzerbyt, V. (2021). When my actions shape your looks: Experience-based properties of approach/avoidance bias the visual representation of others. *Journal of Personality and Social Psychology*, 120(5), 1146-1174. <https://doi.org/10.1037/pspa0000268>
- Rudman, L. A., Dohn, M. C., & Fairchild, K. (2007). Implicit self-esteem compensation: Automatic threat defense. *Journal of Personality and Social Psychology*, 93(5), 798-813. <https://doi.org/10.1037/0022-3514.93.5.798>
- Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology*, 91(6), 995-1008. <https://doi.org/10.1037/0022-3514.91.6.995>
- 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>
- Sassenberg, K., & Ditrich, L. (2019). Research in social psychology changed between 2011 and 2016: Larger sample sizes, more self-report measures, and more online studies.

- Advances in Methods and Practices in Psychological Science*, 2(2), 107-114.
<https://doi.org/10.1177/2515245919838781>
- Sawilowsky, S. S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*, 8(2), 597-599. <https://doi.org/10.22237/jmasm/1257035100>
- Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology*, 4(3), 219-247.
<https://doi.org/10.1037/0278-6133.4.3.219>
- Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47(5), 609-612.
<https://doi.org/10.1016/j.jrp.2013.05.009>
- Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, 51(3), 515-530. <https://doi.org/10.1037/0022-3514.51.3.515>
- Seibt, B., Neumann, R., Nussinson, R., & Strack, F. (2008). Movement direction or change in distance? Self- and object-related approach–avoidance motions. *Journal of Experimental Social Psychology*, 44(3), 713-720.
<https://doi.org/10.1016/j.jesp.2007.04.013>
- Sherman, B. J., Baker, N. L., Squeglia, L. M., & McRae-Clark, A. L. (2018). Approach bias modification for cannabis use disorder: A proof-of-principle study. *Journal of Substance Abuse Treatment*. <https://doi.org/10.1016/j.jsat.2018.01.012>
- Sherman, J. W. (2008). Controlled influences on implicit measures: Confronting the myth of process-purity and taming the cognitive monster. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures* (p. 391-426). Psychology Press.

- Sherman, J. W., Gawronski, B., & Trope, Y. (2014). Two of what. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (p. 3-19). Guildford Press.
- Sigall, H., & Page, R. (1971). Current stereotypes: A little fading, a little faking. *Journal of Personality and Social Psychology*, 18(2), 247-255. <https://doi.org/10.1037/h0030839>
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359-1366. <https://doi.org/10.1177/0956797611417632>
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2013). Life after p-hacking. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2205186>
- Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, 143(2), 534-547. <https://doi.org/10.1037/a0033242>
- Skinner, B. F. (1953). *Science and human behavior*. Macmillan.
- Smith, C. T., Calanchini, J., Hughes, S., Van Dessel, P., & De Houwer, J. (2019). The impact of instruction- and experience-based evaluative learning on IAT performance: A Quad model perspective. *Cognition and Emotion*, 1-21. <https://doi.org/10.1080/02699931.2019.1592118>
- Smith, E. 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

- Solarz, A. K. (1960). Latency of instrumental responses as a function of compatibility with the meaning of eliciting verbal signs. *Journal of Experimental Psychology*, 59(4), 239-245. <https://doi.org/10.1037/h0047274>
- Spatola, N., & Wudarczyk, O. A. (2020). Implicit attitudes towards robots predict explicit attitudes, semantic distance between robots and humans, anthropomorphism, and prosocial behavior: From attitudes to human–robot interaction. *International Journal of Social Robotics*. <https://doi.org/10.1007/s12369-020-00701-5>
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35(1), 4-28. <https://doi.org/10.1006/jesp.1998.1373>
- Stewart, B. D., & Payne, B. K. (2008). Bringing automatic stereotyping under control: Implementation intentions as efficient means of thought control. *Personality and Social Psychology Bulletin*, 34(10), 1332-1345. <https://doi.org/10.1177/0146167208321269>
- Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42(4), 1096-1104. <https://doi.org/10.3758/BRM.42.4.1096>
- Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44(1), 24-31. <https://doi.org/10.1177/0098628316677643>
- 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
- Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, 33(4), 529-554. <https://doi.org/10.1086/214483>

- Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, 76(3), 282-299. <https://doi.org/10.1037/h0027242>
- Unkelbach, C., & Högden, F. (2019). Why Does George Clooney Make Coffee Sexy? The Case for Attribute Conditioning. *Current Directions in Psychological Science*, 096372141985935. <https://doi.org/10.1177/0963721419859354>
- Van Dessel, P., De Houwer, J., & Gast, A. (2016). Approach–avoidance training effects are moderated by awareness of stimulus–action contingencies. *Personality and Social Psychology Bulletin*, 42(1), 81-93. <https://doi.org/10.1177/0146167215615335>
- Van Dessel, P., De Houwer, J., Gast, A., Roets, A., & Smith, C. T. (2020). On the effectiveness of approach-avoidance instructions and training for changing evaluations of social groups. *Journal of Personality and Social Psychology*. <https://doi.org/10.1037/pspa0000189>
- Van Dessel, P., De Houwer, J., Gast, A., Smith, C. T., & De Schryver, M. (2016). Instructing implicit processes: When instructions to approach or avoid influence implicit but not explicit evaluation. *Journal of Experimental Social Psychology*, 63, 1-9. <https://doi.org/10.1016/j.jesp.2015.11.002>
- Van Dessel, P., De Houwer, J., Gast, A., & Tucker Smith, C. (2015). instruction-based approach-avoidance effects: changing stimulus evaluation via the mere instruction to approach or avoid stimuli. *Experimental Psychology*, 62(3), 161-169. <https://doi.org/10.1027/1618-3169/a000282>
- Van Dessel, P., De Houwer, J., Roets, A., & Gast, A. (2016). Failures to change stimulus evaluations by means of subliminal approach and avoidance training. *Journal of Personality and Social Psychology*, 110(1), e1-e15. <https://doi.org/10.1037/pspa0000039>

- Van Dessel, P., Gawronski, B., Smith, C. T., & De Houwer, J. (2017). Mechanisms underlying approach-avoidance instruction effects on implicit evaluation: Results of a preregistered adversarial collaboration. *Journal of Experimental Social Psychology*, 69, 23-32. <https://doi.org/10.1016/j.jesp.2016.10.004>
- Van Dessel, P., Hughes, S., & De Houwer, J. (2018a). How do actions influence attitudes? An inferential account of the impact of action performance on stimulus evaluation. *Personality and Social Psychology Review*, 108886831879573. <https://doi.org/10.1177/1088868318795730>
- Van Dessel, P., Hughes, S., & De Houwer, J. (2018b). Consequence-based approach-avoidance training: A new and improved method for changing behavior. *Psychological Science*, 29(12), 1899-1910. <https://doi.org/10.1177/0956797618796478>
- Van Dessel, P., Mertens, G., Smith, C. T., & De Houwer, J. (2017). The mere exposure instruction effect: Mere exposure instructions influence liking. *Experimental Psychology*, 64(5), 299-314. <https://doi.org/10.1027/1618-3169/a000376>
- Vandenbosch, K., & De Houwer, J. (2011). Failures to induce implicit evaluations by means of approach-avoid training. *Cognition & Emotion*, 25(7), 1311-1330. <https://doi.org/10.1080/02699931.2011.596819>
- Vazire, S. (2018). Implications of the credibility revolution for productivity, creativity, and progress. *Perspectives on Psychological Science*, 13(4), 411-417. <https://doi.org/10.1177/1745691617751884>
- Versace, R., Vallet, G. T., Riou, B., Lesourd, M., Labeye, É., & Brunel, L. (2014). Act-In: An integrated view of memory mechanisms. *Journal of Cognitive Psychology*, 26(3), 280-306. <https://doi.org/10.1080/20445911.2014.892113>

- Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the **metafor** Package. *Journal of Statistical Software*, 36(3). <https://doi.org/10.18637/jss.v036.i03>
- Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, 14(5), 779-804. <https://doi.org/10.3758/BF03194105>
- Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage–Dickey method. *Cognitive Psychology*, 60(3), 158-189.
<https://doi.org/10.1016/j.cogpsych.2009.12.001>
- Wang, M.-T., & Degol, J. L. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. *Educational Psychology Review*, 29(1), 119-140.
<https://doi.org/10.1007/s10648-015-9355-x>
- Warrington, E. K., & Weiskrantz, L. (1968). New method of testing long-term retention with special reference to amnesic patients. *Nature*, 217(5132), 972-974.
<https://doi.org/10.1038/217972a0>
- Wellek, S. (2010). *Testing statistical hypotheses of equivalence and noninferiority* (2nd ed). CRC Press.
- Wentura, D., & Degner, J. (2010). A practical guide to sequential priming and related tasks. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (p. 95-116). The Guilford Press.
- Wicker, A. W. (1969). Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, 25(4), 41-78.
<https://doi.org/10.1111/j.1540-4560.1969.tb00619.x>
- Wiers, C. E., Kühn, S., Javadi, A. H., Korucuoglu, O., Wiers, R. W., Walter, H., Gallinat, J., & Bermpohl, F. (2013). Automatic approach bias towards smoking cues is present in

- smokers but not in ex-smokers. *Psychopharmacology*, 229(1), 187-197.
<https://doi.org/10.1007/s00213-013-3098-5>
- Wiers, R. W., Eberl, C., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2011). Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychological Science*, 22(4), 490-497.
<https://doi.org/10.1177/0956797611400615>
- Wiers, R. W., Gladwin, T. E., Hofmann, W., Salemink, E., & Ridderinkhof, K. R. (2013). Cognitive bias modification and cognitive control training in addiction and related psychopathology: Mechanisms, clinical perspectives, and ways forward. *Clinical Psychological Science*, 1(2), 192-212. <https://doi.org/10.1177/2167702612466547>
- Williamson, V. (2016). On the ethics of crowdsourced research. *PS: Political Science & Politics*, 49(01), 77-81. <https://doi.org/10.1017/S104909651500116X>
- Wittekind, C. E., Feist, A., Schneider, B. C., Moritz, S., & Fritzsche, A. (2015). The approach-avoidance task as an online intervention in cigarette smoking: A pilot study. *Journal of Behavior Therapy and Experimental Psychiatry*, 46, 115-120.
<https://doi.org/10.1016/j.jbtep.2014.08.006>
- Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology*, 81(5), 815-827. <https://doi.org/10.1037/0022-3514.81.5.815>
- Wolsiefer, K., Westfall, J., & Judd, C. M. (2017). Modeling stimulus variation in three common implicit attitude tasks. *Behavior Research Methods*, 49(4), 1193-1209.
<https://doi.org/10.3758/s13428-016-0779-0>
- Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology*, 10(2), 109-120. [https://doi.org/10.1016/0022-1031\(74\)90059-6](https://doi.org/10.1016/0022-1031(74)90059-6)

- Woud, M. L., Becker, E. S., & Rinck, M. (2008). Implicit evaluation bias induced by approach and avoidance. *Cognition & Emotion*, 22(6), 1187-1197.
<https://doi.org/10.1080/02699930801906843>
- Woud, M. L., Becker, E. S., & Rinck, M. (2011). Induction of implicit evaluation biases by approach–avoidance training: A commentary on Vandenbosch and De Houwer (this issue). *Cognition and Emotion*, 25(7), 1331-1338.
<https://doi.org/10.1080/02699931.2011.628300>
- Woud, M. L., Maas, J., Becker, E. S., & Rinck, M. (2013). Make the manikin move: Symbolic approach–avoidance responses affect implicit and explicit face evaluations. *Journal of Cognitive Psychology*, 25(6), 738-744.
<https://doi.org/10.1080/20445911.2013.817413>
- Yzerbyt, V. Y., Leyens, J.-P., & Corneille, O. (1998). Social judgeability and the bogus pipeline: The role of naive theories of judgment in impression formation. *Social Cognition*, 16(1), 56-77. <https://doi.org/10.1521/soco.1998.16.1.56>
- Yzerbyt, V. Y., Schadron, G., Leyens, J.-P., & Rocher, S. (1994). Social judgeability: The impact of meta-informational cues on the use of stereotypes. *Journal of Personality and Social Psychology*, 66(1), 48-55. <https://doi.org/10.1037/0022-3514.66.1.48>
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9(2, Pt.2), 1-27. <https://doi.org/10.1037/h0025848>
- Zanon, R., De Houwer, J., Gast, A., & Smith, C. T. (2014). When does relational information influence evaluative conditioning? *Quarterly Journal of Experimental Psychology*, 67(11), 2105-2122. <https://doi.org/10.1080/17470218.2014.907324>

APPENDICES

Appendix 1
Acronym Table

Acronym	Definition
AA	Approach-Avoidance
APE	Associative-Propositional Evaluation
IAT	Implicit Association Test
NHST	Null Hypothesis Significance Testing
SC-IAT	Single Category Implicit Association Test
RT	Response Time
SEM	System of Evaluations Model
SESOI	Smallest Effect Size Of Interest
SS	Sample Size
STEM	Science, Technology, Engineering, and Mathematics
VAAST	Visual Approach-Avoidance by the Self Task

Appendix 2

Replication Recipe for the Replication of Kawakami et al. (2008; Brandt, et al., 2014)

The Nature of the Effect

1. Verbal description of the effect I am trying to replicate:

We want to replicate the effect approach and avoidance training have on indirect evaluations. Specifically, we want to see if repeatedly approaching math-related material (instead of avoiding it) increase implicit identification (as measured with an identification IAT). The effect is originally described in Kawakami et al. (2008): Participants who had to approach math and avoiding arts (instead of the opposite) were faster in the IAT block where they categorized math and the self with the same response key.

2. It is important to replicate this effect because:

This effect offers a solution to the gender gap in STEM problem. Nevertheless, it suffers from several problems: low sample size, lack of preregistered replication, and ambiguous implementation of approach and avoidance movements.

3. The effect size of the effect I am trying to replicate is:

Meta-analytic effect size is $d_s = 0.64$ (Kawakami et al., 2008; Exp. 1 and Exp. 2).

4. The confidence interval of the original effect is:

$CI_{95\%}$ for the meta-analytic effect size is [0.17; 1.11]

5. The sample size of the original effect is:

Exp. 1: 24 participants (total: 44)

Exp. 2: 56 participants

6. Where was the original study conducted? (e.g., lab, in the field, online)

Original experiment was conducted in lab.

7. What country/region was the original study conducted in?

Original experiments took place in Canada.

8. What kind of sample did the original study use? (e.g., student, Mturk, representative)

Original sample was constituted of undergraduate students. Note that they were all female participants who reported not to like math.

9. Was the original study conducted with paper-and-pencil surveys, on a computer, or something else?

Original experiments were computerized.

Designing the Replication Study

10. Are the original materials for the study available from the author?

No.

a. If not, are the original materials for the study available elsewhere (e.g., previously published scales)?

Identification IAT is extensively described in other papers (e.g., Nosek et al., 2003)

b. If the original materials are not available from the author or elsewhere, how were the materials created for the replication attempt?

Materials were created using Method section of Kawakami et al.'s (2008) manuscript. We had to reproduce the identification IAT, and to build a data set of 24 images related to math and 24 images related to art. We used Wikimedia to create this data set. Material is available on the OSF project of this manuscript.

11. I know that assumptions (e.g., about the meaning of the stimuli) in the original study will also hold in my replication because:

- VAAST-base AAT can form evaluative response (Experiment 1a).
- Stimuli were picked to represent math and art category

12. Location of the experimenter during data collection:

Experimenter led participant to an individual cubicle and then leave.

13. Experimenter knowledge of participant experimental condition:

No. The computer program automatically picked a condition for the participant without Experimenter knowledge.

14. Experimenter knowledge of overall hypotheses:

Yes.

15. My target sample size is:

200 participants.

16. The rationale for my sample size is:

200 participants sample size gives 80% statistical power to detect a small-to-medium effect size ($1 - \beta = .80$ for $d_s = 0.40$). The 60% CI lower bound of the effect size one can estimate from Kawakami et al. (2008) 's experiments corresponds to a an effect size of $d_s = 0.44$.

Documenting Differences between the Original and Replication Study

17. The similarities/differences in the instructions are: Close.

We could not gain access to original experience. We rely on the Method section of Kawakami et al. (2008)'s manuscript to recreate the instructions.

18. The similarities/differences in the measures are: Exact.

We used material by Nosek et al. (2002) to recreate the Identification IAT.

19. The similarities/differences in the stimuli are: Close.

We could not gain access to material used by Kawakami et al. (2008). We build our own image dataset using images from Wikimedia. We translated words used for the identification IAT in French using Nosek et al. (2002) material.

20. The similarities/differences in the procedure are: Different.

The procedure for the experiment is the same except for the AAT implementation. Instead of using the Joystick task, we used an implementation inspired from Rougier et al. (2018). This task is more valid than the Joystick task and its effect is not smaller than the Joystick task.

21. The similarities/differences in the location (e.g., lab vs. online; alone vs. in groups) are: Exact.

Both original experiment and replication took place in lab.

22. The similarities/differences in remuneration are: Close.

Undergraduate students could be either paid 10€ or with course credit. Non undergraduate participants were paid 10€ (see Population).

23. The similarities/differences between participant populations are: Close.

Every participant in Kawakami et al. (2008) original experiments were undergraduate female participant with low identification to math. We kept the same population but also recruited non-undergraduate participants with low identification to math. There is no reason why effect might differ between this two groups of participants.

24. What differences between the original study and your study might be expected to influence the size and/or direction of the effect?

Using a VAAST-based implementation (Rougier et al., 2018) instead a Joystick task one (Kawakami et al., 2008) could increase the effect size. Nevertheless, we are not primarily interested in the difference.

25. I have taken the following steps to test whether the differences listed in #24 will influence the outcome of my replication attempt:

We conducted a pilot experiment showing that in the context of evaluation formation, the Joystick task is not better than the VAAST.

Analysis and Replication Evaluation

26. My exclusion criteria are (e.g., handling outliers, removing participants from analysis):

1. Participants who did not complete the whole study.
2. Statistical outlier according to Judd, McClelland, and Ryan criteria (Lever, Cook's distance, studentized deleted residuals, 2009).

27. My analysis plan is (justify differences from the original):

t test comparing IAT difference score between approach math condition and avoid math condition. IAT difference score will be computed with the same procedure Kawakami et al. (2008) used. This is the exact same test Kawakami et al. (2008) to test the hypothesis that AAT could increase math implicit self-identification. This exact analysis was preregistered.

28. A successful replication is defined as:

Estimation of the effect size lies within the effect size computed from Kawakami et al. (2008) data ($CI_{95\%}[0.17; 1.11]$) and AAT effect is significant.

Registering the Replication Attempt

29. The finalized materials, procedures, analysis plan etc of the replication are registered here:

osf.io/5yvdy (pre-registration)

Reporting the Replication

30. The effect size of the replication is:

$$d_s = 0.30$$

31. The confidence interval of the replication effect size is:

$$d_s \text{ } CI_{95\%} [0.02; 0.57]$$

32. The replication effect size is not significantly different from the original effect size.

33. I judge the replication to be a success because:

Approach and avoidance training effect on Identification IAT is significant, estimated effect size falls within the Kawakami et al. (2008) $CI_{95\%}$ of the meta-analytic effect size.

34. Interested experts can obtain my data and syntax here:

<https://osf.io/pcea3/>

35. All of the analyses were reported in the report or are available here:

<https://osf.io/amq99/>

36. The limitations of my replication study are:

We consider this replication successful given its original aims. The increase in sample size and the meta-analysis provides a more precise estimation of the effect of AAT on math identification, $d_s = 0.40$, $CI_{95\%}$ [0.14; 0.65]

Power analyses reveal that this study is adequately powered given Kawakami et al.'s (2008) effect size. The replication was preregistered, and we do not report deviation from the preregistration. We used an improved paradigm that we validate in a previous study.

Appendix 3

Material used in Chapter 6

Table 1

Stimuli Used for the Approach and Avoidance Training in Experiment 1a and 1b

Luupite category	Maasolup
	Neenolup
	Omeelup
	Tuuralup
	Wenaalup
Niffite category	Cellanif
	Eskannif
	Lebunnif
	Otrannif
	Zallunif

Note. Stimuli were retrieved from Gregg et al. (2006).

Table 2*Stimuli Used in the IAT of Experiment 1a and 1b*

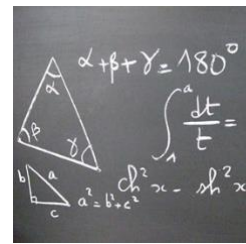
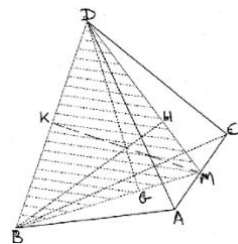
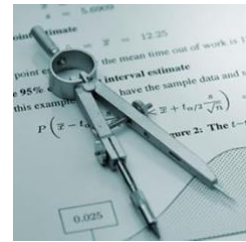
Luupite category	Maasolup	
	Neenolup	
	Omeelup	
	Tuuralup	
	Wenaalup	
Niffite category	Cellanif	
	Eskannif	
	Lebunnif	
	Otrannif	
	Zallunif	
Self category	Je	(I)
	Mien	(Mine)
	Moi	(Myself)
	Mon	(My)
Other category	Eux	(Them)
	Eux-mêmes	(Themselves)
	Ils	(They)
	Leur	(Their)

Note. Words in parenthesis are English translation from the French words used in the Experiment 1a. The English words were use in Experiment 1b. Stimuli were retrieved from Gregg et al. (2006) and inspired from Nosek et al. .

Table 3

Example of Stimuli Used for the Approach and Avoidance Training in Experiment 2

Stimuli related to math



Stimuli related to art



Note. The exhaustive set of stimuli contained 48 images (24 related to math and 24 related to art). Images were retrieved from Wikimedia. Full material is available at <https://osf.io/pcea3/>.

Table 4*Stimuli Used in the IAT in Experiment 2*

Math category	Calcul	(Calculus)
	Équation	(Equation)
	Géométrie	(Geometry)
	Statistiques	(Statistics)
Art category	Littérature	(Literature)
	Poésie	(Poetry)
	Symphonie	(Symphony)
	Théâtre	(Theatre)
Self category	Je	(I)
	Mien	(Mine)
	Moi	(Myself)
	Mon	(My)
Other category	Eux	(Them)
	Eux-mêmes	(Themselves)
	Ils	(They)
	Leur	(Their)

Note. Stimuli were inspired from Nosek et al. (2002). Words in parenthesis are English translation from the French words used in the experiment.

Appendix 4

Mathematical models used in Experiments 3, 4, and 5

In experiments 3, 4, and 5, we analyzed data from an AMP using logistic mixed modeling. While non-unusual in the literature, this approach is still relatively new (Wolsiefer et al., 2017). Among the difficulties that exist with a mixed-modeling approach, the algorithms used to estimate the values of the parameters in the mixed model sometimes fail to converge and this risk is even higher in the case of logistic mixed modeling (Eager & Roy, 2017). Accordingly, in Experiments 4 and 5, we were confronted to such convergence problem. To fix these kinds of issues, Bates et al. (2015) recommend adopting a simplified version of the models. This appendix contains the *lme4* syntax for the models that we ended up using.

Experiment 3

```
glmer(amp_response_n ~
      amp_movement_c * cond_training_c * amp_stimulus_c +
      (1 + amp_stimulus_c | participant),
      family = binomial,
      data = dataset)
```

Experiment 4

```
glmer(amp_response_n ~
      amp_movement_c * amp_training_c +
      (1 + amp_movement_c + amp_training_c | participant) +
      (0 + amp_training_c | amp_logo_filename),
      family = binomial,
      data = dataset)
```

Experiment 5

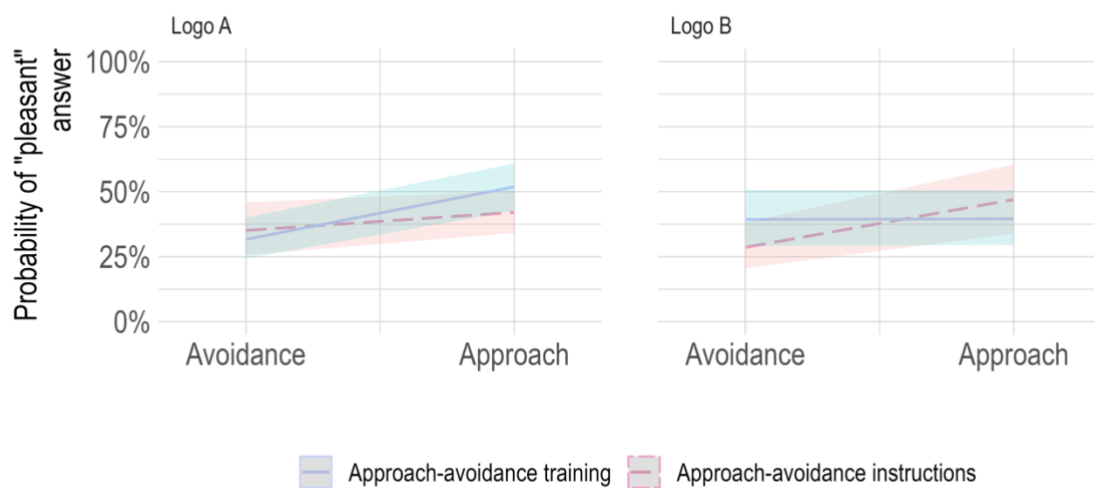
```
glmer(amp_response_n ~
      cond_training_c * amp_movement_c * amp_type_c +
      (1 + amp_movement_c * amp_type_c | participant) +
      (0 + amp_movement_c | amp_prime),
      family = binomial,
      data = dataset)
```


Appendix 5

Results of Experiment 4

Figure 1

Probability of a “Pleasant” Answer During the AMP according to the Logo Used as a Prime, the Experience with the Prime, and the Movement the Prime was Associated With



Note. Confidence interval represents the 95% CI of prediction at the population level based on the statistical model.

Appendix 6

Stimuli Used in Experiment 6 and 7

Table 1

Stimuli Used in Experiment 6 and 7

Luupite category	Maasolup
	Neenolup
	Omeelup
	Tuuralup
	Wenaalup
Niffite category	Cellanif
	Eskannif
	Lebunnif
	Otrannif
	Zallunif
Self category	I
	Mine
	Myself
	My
Others category	Them
	Themselves
	They
	Their

Note. Stimuli were retrieved from Gregg et al. (2006) and inspired from Nosek et al. (2002).

Appendix 7

Verbatim from Experiment 8

In the fictitious experiment we are going to present you, we are interested in creating implicit attitudes toward groups of people. Implicit attitudes refer to the affective spontaneous reaction something can trigger. As an example, on the one hand, a piece of cake could probably trigger a positive gut feeling: a positive implicit attitude. On the other hand, a snake would probably trigger a negative gut feeling: a negative implicit attitude.

In the fictitious studies, researchers are aiming to form implicit attitudes toward two different social groups: the Luupites and the Niffites. Each group is very different from each other, and every person from the two groups can be easily recognized for the group they belong to.

- [Page break] -

- [Approach-Avoidance instructions first condition] –

To create attitudes toward these groups, the researchers are planning on using either one of two methods, and each method will be tested in one of two conditions. If you were a participant, you would either be assigned to the first method, the “instructions” condition, or to the second method, the “training” condition. We will now present you these two conditions. Be very careful because several questions will be asked later and the descriptions of the conditions will not be presented again.

- [Page break] -

In the “instructions” condition, you would receive instructions indicating that in a task that will follow, you will have to approach members of the Niffite group and avoid members of the Luupite group. The researchers would explain to you that you will receive complementary information on how to approach and how to avoid the group members later on, but first you would have to go through another unrelated task. Thus, you would not actually perform the approach and avoidance movements toward the group members. Instead, after receiving just the instructions for the approach/avoidance task, you would complete an Implicit Association

Test which measures your implicit attitudes toward the Luupite and the Niffite, that is, the spontaneous affective reactions these groups evoke.

In the “training” condition, you would receive instructions indicating that in a task that will follow, you will have to approach members of the Niffite group and avoid members of the Luupite group. The researchers would then tell you how to proceed: Group members’ names will appear in a virtual environment like in a video game representing a street. In this environment, you will have to approach and avoid members as previously instructed, by using the arrow keys on your keyboard. You would then actually complete the training which consists of 200 trials; it would last approximately 10 minutes. After this training, you would complete an Implicit Association Test which measures your implicit attitudes toward the Luupite and the Niffite, that is, the spontaneous affective reactions these groups evoke.

Please make sure that you understood the different conditions. The questions we will ask are critical for this research.

- [Approach-Avoidance training first condition] –

To create attitudes toward these groups, the researchers are planning on using either one of two methods, and each method will be tested in one of two conditions. If you were a participant, you would either be assigned to the first method, the “training” condition, or to the second method, the “instruction” condition. We will now present you these two conditions. Be very careful because several questions will be asked later and the descriptions of the conditions will not be presented again.

- [Page break] -

In the “training” condition, you would receive instructions indicating that in a task that will follow, you will have to approach members of the Niffite group and avoid members of the Luupite group. The researchers would then tell you how to proceed: Group members’ names will appear in a virtual environment like in a video game representing a street. In this environment, you will have to approach and avoid members as previously instructed, by using the arrow keys on your keyboard. You would then actually complete the training which

consists of 200 trials; it would last approximately 10 minutes. After this training, you would complete an Implicit Association Test which measures your implicit attitudes toward the Luupite and the Niffite, that is, the spontaneous affective reactions these groups evoke.

In the “instructions” condition, you would receive instructions indicating that in a task that will follow, you will have to approach members of the Niffite group and avoid members of the Luupite group. The researchers would explain to you that you will receive complementary information on how to approach and how to avoid the group members later, but first you would have to go through another unrelated task. Thus, you would not actually perform the approach and avoidance movements toward the group members. Instead, after receiving just the instructions for the approach/avoidance task, you would complete an Implicit Association Test which measures your implicit attitudes toward the Luupite and the Niffite, that is, the spontaneous affective reactions these groups evoke.

Please make sure that you understood the different conditions. The questions we will ask are critical for this research