ISPA

INSTITUTO UNIVERSITÁRIO

CIÊNCIAS PSICOLÓGICAS, SOCIAIS E DA VIDA

THE KNOWN AND UNKNOWN OF SOCIAL FACILITATION ON STEREOTYPING

Pedro Miguel Regueiras Figueira

Tese submetida como requisito parcial para obtenção de

Doutoramento em Psicologia

Área de especialidade Psicologia Social

2020

ISPA

INSTITUTO UNIVERSITÁRIO

CIÊNCIAS PSICOLÓGICAS, SOCIAIS E DA VIDA

THE KNOWN AND UNKNOWN OF SOCIAL FACILITATION ON STEREOTYPING

Pedro Miguel Regueiras Figueira

Tese orientada por

Teresa Garcia-Marques, PhD. (ISPA-Instituto Universitário, Lisboa, Portugal)

e

Jeff Sherman, PhD (University of California, Davis, USA)

Tese submetida como requisito parcial para obtenção do grau de
Doutoramento em Psicologia
Área de especialidade Psicologia Social

2020

Tese apresentada para cumprimento dos requisitos necessários à obtenção do grau Doutor em Psicologia na área de especialização de Psicologia Social realizada sob a orientação de Teresa Garcia-Marques, PhD. e Jeff Sherman PhD. apresentada no ISPA-Instituto Universitário no ano de 2020.

O presente trabalho foi financiado pelo Lisbon PhD in Social Psychology (LISP) um programa da Fundação para a Ciência e Tecnologia (PD/BD/113470/2015).

Ao Sonhador,

que nos sonha…

Agradecimentos

Esta tese foi um caminho, uma peregrinação concluída com ajuda de muitos. Da mesma forma que David Mitchell escreveu "Our lives are not our own. We are bound to others, past and present, and by each crime and every kindness, we birth our future", esta tese também não é minha. Por isso só me resta agradecer a tudo e todos que me ajudaram nesta peregrinação.

Primeiro agradeço as todas as forças invisíveis à minha compreensão que me ajudaram neste caminho, e que o tornaram possível.

Agradeço aos meus criadores humanos. Á minha mãe que sempre me disciplinou e acreditou em mim, e que quando era preciso fazia de tripas coração. Ao meu pai que sempre se sacrificou pelo bem dos que ama. Espero um dia que a vida vos dê leveza que tanto merecem ao terem carregado este vosso filho nesta vida, e neste caminho.

Agradeço ao Amor, à Paixão, à Amizade e seus avatares. Vocês destruíram preconceitos, juízos de valor, aqueceram-me nas noites mais frias, curaram-me tantas feridas, cobriram as minhas fraquezas e mais importante do que tudo, deram-me o alento para continuar. Espero conseguir retribuir-vos tudo que me deram nesta vida, e neste caminho.

Agradeço aos professores. A todos que me acarinharam, que me desafiaram porque sabiam que eu conseguia, que foram pacientes comigo, e que me deram oportunidades de subir esta escada invisível do conhecimento humano. Espero que a minha vida, e este caminho, dignifiquem as vossas lições.

Agradeço aos obstáculos, e seus avatares. Vocês obrigaram-me a fazer escolhas, lidar com emoções, e criar intenções. Sem vocês seria um barco à deriva, sem capitão.

Finalmente agradeço aos mestres, avatares do Conhecimento. Neste caminho vocês transformaram montanhas em planícies, fogo em água, sofrimento em discernimento.

Grato.

Palavras-chave:
Facilitação Social; Estereótipos; Modelos Dualistas

Key words:
Social Facilitation; Stereotypes; Dual Process Theory

Categorias de Classificação da tese
(Utilizar PsycINFO Classification Categories and Codes):
Exemplo:
2300 Human Experimental Psychology
    2340 Cognitive Processes

3000 Social Psychology
    3040 Social Perception and Cognition

RESUMO

Desde Triplett (1898) que a psicologia experimental tem vindo a explorar de que forma a presença de outros (em contraste com estar sozinho) afeta o nosso comportamento e mente. Este efeito é atualmente conhecido como Facilitação Social. Embora muitos avanços tenham sido feitos nesta área, pouco é conhecido sobre de que forma a Facilitação Social afeta a estereotipização (uso dos estereótipos) e os seus mecanismos.

Até à data apenas dois artigos abordaram esta temática, sendo que estes apresentam conclusões opostas. Enquanto Lambert et al. (2003) sugerem mais estereotipização na presença de outros, Castelli e Tomelleri (2008) sugerem menos estereotipização na presença de outros. De forma a abordar esta incongruência na literatura, replicamos conceptualmente a experiência principal de cada um dos artigos. Os nossos resultados não replicaram nenhuma das conclusões dos artigos originais. Não encontrámos um efeito de Facilitação Social claro sobre a estereotipização quando seguimos a metodologia de Lambert et al. (2003). Já com a replicação de Castelli e Tomelleri (2008) encontrámos evidência de Facilitação Social sobre estereotipização, mas na direção oposta da apresentada no artigo original, ou seja, maior estereotipização na presença de outros.

Ao longo desta tese foram desenvolvidos estudos e analisados resultados de forma cuidadosa com o objetivo de se obter uma melhor compreensão sobre os efeitos em estudo. Já que os dados sugeriram que os efeitos em estudo poderiam se estar a propagar pelos tempos de reação e taxas de erro, analisámos os nossos dados (e os dados de uma nova experiência) através do Modelo de Difusão. Esta técnica permite juntar tempos de reação e taxas de acerto/erro no conjunto de parâmetros. Esta análise permitiu alcançar informação relevante para futuros estudos que abordam os efeitos de estereotipização como também os efeitos da presença ou isolamento de outros. Em primeiro lugar, foram detetados efeitos de Facilitação Social sobre estereotipização, visto que a estereotipização foi maior na presença de outros no que no isolamento. Em segundo lugar, confirmamos que os tipos de condições sociais são relevantes para o efeito, uma vez que a nossa operacionalização (co-ação versus isolamento-sem a presença do experimentador), originou resultados divergentes dos estudos originais. Em terceiro lugar, verificou-se que os efeitos de Facilitação Social sobre a estereotipização dependem do tipo de tarefa usada para medir a estereotipização. Estes efeitos foram mais subtis na Weapon Identification Task (WIT; Payne, 2001) e mais claros na Implicit Association Task (IAT; Greenwald,

McGhee & Schwartz, 1998). Nós argumentamos que isto ocorre porque os efeitos de Facilitação Social ocorrem através de mecanismos que são representados de forma diferente em cada uma das tarefas. Em quarto lugar, a nossa condição de isolamento desafia os resultados que são tipicamente obtidos na WIT. Desta forma, os nossos dados são interpretados como evidência que diferentes condições sociais podem levar as pessoas a lidarem com a tarefa de forma diferente. Além disso, os nossos dados de forma geral evidenciam mais viés estereotípico e menos controlo sobre a ativação do estereótipo na presença de outros. Interpretamos estes dados como estando a corroborar os estudos anteriores que enfatizam a presença de outros como um contexto que cria sobrecarga cognitiva (Baron, 1986) e que de alguma forma reduz a capacidade de usar mecanismos de controlo de forma eficiente (Wagstaff, et al., 2008).

ABSTRACT

Since Triplett (1898), that experimental psychology has explored how the presence of others (vs being alone) affects our behavior and mind. This effect is nowadays known as Social Facilitation. Although, many advances have been made in this area, little is known about how Social Facilitation affects stereotyping. As such, this thesis investigates how the presence of other persons (vs being alone) affects our stereotyping and its mechanisms.

Until the present date, only two papers have addressed this theme reaching opposite conclusions. While Lambert, et al. (2003) suggest more stereotyping in the presence of others, Castelli and Tomelleri (2008) suggest less stereotyping in the presence of others. To approach this incongruency in the literature, we conceptually replicated the main experiment of each of these papers. Our results did not replicate any of those papers. We did not find a clear Social Facilitation effect over stereotyping when following Lambert, et al's. (2003) methodology. Moreover, when replicating Castelli and Tomelleri (2008), we found evidence of a Social Facilitation effect over stereotyping, but now in the opposite direction of the original study, showing more stereotyping in presence of others.

Throughout this thesis, we developed studies and carefully analyzed the results aiming for better understanding the effects. Since data suggested that the evidence could be spread over reaction times (RTs) or error rates, we analyzed our data (and data from a new experiment) by using the Diffusion Model. This technique allows assembling RTs and accuracy data into a set of parameters. This analysis was highly fruitful adding relevant information for future empirical approaches to stereotyping effects in the presence and isolation from others. First, Social Facilitation effects over stereotyping occur, because those in presence of others have higher stereotyping than those in isolation. Second, as already stated in the Social Facilitation literature, the type of Social Condition matters. Because we operationalized social presence as co-action and isolation having no presence of the experimenter, results diverge from the original studies. Third, Social Facilitation effects over stereotyping depend on the type the task used to measure stereotyping. These effects were more subtle in the Weapon Identification Task (WIT; Payne, 2001) and clearer in the race Implicit Association Task (IAT; Greenwald, McGhee & Schwartz, 1998). We argue that this occurs because Social Facilitation effects occur through a mechanism that is differently represented in those tasks. Fourth, our isolation condition challenges the typical results obtained in the WIT. As such our data is

interpreted as evidence that different social conditions can lead people to cope differently with this task. Moreover, our data in general shows evidence of more stereotype bias and less control over stereotype activation in presence of others. We interpret this data as corroborating evidence of previous claims that the presence of others creates an overload context (Baron, 1986) and that it somehow reduces our capability of exerting an efficient control mechanism (Wagstaff, et al., 2008).

.

Index

Chapter I Social Facilitation 3

Social facilitation: The social psychology kick-off 3

Social Facilitation Theories: The right moderator, the right mechanisms? 7

Old problems, new problems: The path for clarification. 18

Attention, Distraction, and Executive Control, plus Context-Sensitivity effects: A new SF era. 25

Chapter II Stereotyping and effects of Social Facilitation 40

Stereotypes as structures of knowledge, and their activation 40

Stereotype activation 43

Stereotype expression: Inevitable or controllable? 47

How stereotype activation and control interplay has been studied. 49

Research on Social Facilitation effects on stereotyping 56

Chapter III The effects of Social facilitation on Stereotyping: When replication reveals the unexpected 61

Experiment 1- Weapon Identification Task (WIT) with actual SF conditions 64

Experiment 2 - Implicit Association Test with SF conditions 76

Chapter IV A Diffusion Model approach 87

The Diffusion Model-Computation of mechanisms through responses and reactions times. 87

Diffusion Model on Stereotyping 92

Diffusion model in IAT and WIT 98

How to use the Diffusion model to model WIT experimental data 99

Experiment 1 in light of the Diffusion Model 101

Experiment 3- Implicit Association Test with restriction of time 112

Chapter V General Discussion 127

Further Questions 132

Conclusion 136

References 138

Attachments 160

Table and Figure Index

Figure 1- Weapon Identification Trials 66

Table 1- Self-Report Measures......67

Figure 2-Error Rates (Experiment 1)......70

Table 2- PDP Estimates(Experiment 1)......25

Figure 3-Correct Reaction Times (Experiment 1) 73

Figure 4-CDF Plots (Experiment 1) 74

Figure 5- IAT Trials 78

Figure 6- IAT Trials 79

Figure 7- *d600* (Experiment 2)......81

Figure 8- Error Rates (Experiment 2) 81

Table 3- PDP Estimates (Experiment 2) 82

Figure 9- Correct Reaction Times (Experiment 2)......83

Figure 10-CDF Plots (Experiment 2) 84

Figure 11- Diffusion Model Illustration......89

Table 4- Diffusion Model Parameter interpretation (WIT) 100

Figure 12- Drift-Rates for WIT (Congruent vs Incongruent)...... 102

Figure 13- Relative Starting-Points for WIT (Congruent vs Incongruent) 103

Figure 14- Threshold Separation (Congruent vs Incongruent)...... 104

Figure 15-Nondecision Times (Congruent vs Incongruent) 105

Figure 16- Drift-Rates for WIT (Gun vs Tool)...... 107

Figure 17- Relative Starting-Points for WIT (Gun vs Tool) 108

Figure 18- Threshold Separation for WIT (Gun vs Tool)...... 109

Figure 19- Nondecision Times for WIT (Gun vs Tool) 110

Table 5- Diffusion Model Parameter interpretation (IAT) ... 113

Figure 20- *d600* (Experiment 3) ... 117

Figure 21- Error Rates (Experiment 3) ... 118

Figure 22- Correct Reaction Times (Experiment 3) ... 118

Table 6- PDP Estimates (Experiment 3) ... 119

Figure 23- CDF Plots RTs (Experiment 3) ... 120

Figure 24- CDF Plots ERs (Experiment 3) ... 121

Figure 25- Drift-Rates for IAT ... 122

Figure 26- Relative Starting-Points for IAT ... 123

Figure 27- Threshold Separation for IAT ... 124

Figure 28- Nondecision Times for IAT ... 125

Introduction

Although humans are social beings, that walk through many and different social contexts, we still may find contexts where people isolate from each other. In just one day, as an individual we walk in many different scenarios where we are surrounded by others (except in epidemic crisis scenarios), such as walking on the street, inside a bus, working near co-workers, etc. At same time, we also go into different scenarios where we are alone. For instance, being alone at home, having meal alone, going to bathroom, driving a car, etc. When social researchers ask if these different context matter, they ask the most fundamental question regarding our social features. They ask about the variation in our behaviors and minds occurring between a context where we are in the Presence of Others (PO) versus in No Presence of Others (NPO).

Research has positively answered this question: the presence of others changes our behavior. This effect was initially named as Social Facilitation (SF; Allport 1920). Since Triplett (1898), who was the first to empirical test these ideas, social psychology has examined how PO affects our behavior. However, only more recently, research has asked the question of the effects on cognition and in this regard, there is still a long debate to be developed about how SF effects impact cognitive mechanisms.

In this thesis, we enter in the debate by exploring SF effect on stereotyping. We thought that the crossover between the two fields (SF and Stereotyping) is beneficial, since the stereotyping literature uses and provides many different tasks, estimation techniques, and is anchored in well-defined cognitive mechanisms. Hereby, offering an ideal setting to explore possible cognitive mechanism that sustain SF effects. As such, this thesis aims to explore how PO and NPO modulates our minds, favoring or helping to prevent stereotyping effects.

In the subsequent chapters, we will review the literature both on SF and stereotyping, and present a set of empirical studies to clarify how SF effects can occur on stereotyping.

In Chapter I, we review SF effects through a century of research. Within this review, we focus the most important discovers, the main theories and their problems. We will contrast alternative hypothetical mechanisms that promote SF effects and define SF effects, while distinguishing them from other effects that encompasses the presence of others (e.g., evaluation, social comparison, accountability…).

In Chapter II, we define stereotyping and strategically focus on research that revealed cognitive stereotyping mechanisms. We define what is a Stereotype, and how it influences our perception, judgment and cognition. In our review, we highlight the features of stereotyping that can be affected by SF manipulations. Importantly, here we focus on the two empirical studies (Lambert, et al., 2003; Castelli & Tomelleri, 2008) that claimed to have already addressed SF effects on stereotyping.

In Chapter III, we put together the information gathered from our review and claim that in order to study social facilitation effects, one and the same experimental paradigm has to be used: one that compares real isolated conditions (with no presence of the experimenter), with the presence of others that are not attended to their task (being simple co-actors). Using this paradigm, we propose to conceptually replicate Lambert, et al. (2003) and Castelli and Tomelleri (2008) in two experiments, addressing the hypothesis that their incompatible findings may be explained by their different operationalization of social presence conditions). From a critical discussion and analysis of our results, we raise new questions that led us to follow a different route to explore SF on stereotyping.

In Chapter IV, we define an alternative method that can be implemented to analyze our data. We define the Diffusion Model (DM) as a technical model that allows us to better understand our data. In addition, we develop a new experiment to further test the effects with this new methodology.

In Chapter V, we offer a general discussion based on the previous literature and our data. Here we raise new questions and offer some conclusions that challenges the future of the two fields (SF and stereotyping).

Chapter I Social Facilitation

In this chapter, the concept of SF and the major approaches that have been and are being taken to the study of this phenomenon are reviewed. The first part of the chapter takes a historical look at the concept and outlines how a definition of the Social Facilitation effect and alternative explanations of its occurrence were developed. The second part of the chapter argues for a stricter definition of SF effects, and summarizes the major findings that inform its cognitive features.

Social facilitation: The social psychology kick-off

SF is a sub-field within social psychology that explores the effects of social presence, both physical and imagined, on task performance (Aiello & Douthit, 2001). Despite SF being a minor field inside a major field (i.e., social psychology), it is neither strange nor coincidental that both fields share very similar definitions: "Social psychology is the scientific study of how people's thoughts, feelings, and behaviors are influenced by the actual, imagined, or implied presence of others" (Allport, 1985, p. 5).

The truth about this subtle, and sometimes forgotten, resemblance seems to lie in the fact that social psychology as an experimental and scientific field started with the publication of Triplett's (1898) work. His work can be viewed as the social psychology kick-off, seeding the field for more than a century with amazing discoveries that have led to a better understanding of what it is to be Human in social contexts. At the time, Triplett's concern was to discover how social contexts influence human performance. In his first study, he analyzed the time records of professional cyclists when cycling with other cyclists versus cycling alone against the clock. The analysis revealed that, on average, cyclists are faster when competing with others, that is, in the presence of others (PO), than when competing alone against the clock. Triplett set up a second study, this time with children playing a fishing rod game, in which the aim is to spin a fishing line as fast as possible. The children were split into two groups; in one group, two co-present children did the task at the same time, while in another group, each child did the task alone. The results were similar to the first study: faster performance in PO. For Triplett, this result indicated that PO somehow stimulated a competitive instinct, which lead to a "greater concentration of energy" (Triplett, 1898, p. 526), and thus motivation for faster performances in the assigned tasks. This concept of energy eventually came to denominate arousal, which developed into a key feature in many future SF theories.

About 20 years later, Allport (1920) demonstrated and tested the assumption that behavior in response to a non-social stimulus is modified by the presence and actions of other people. Today, this is accepted as fact; however, until Allport, it had not been tested. Allport's proposition was different and more specific than Triplett's, as he aimed to test SF effects without the possibility of performance comparison—the tasks used by Triplett (1898) were competitive in nature, which encouraged individuals to be faster than another person. Allport's challenge was to create an experimental paradigm that would eliminate the effects of competition, and that would rely on tasks that were impossible to be monitored by others, hence eliminating performance comparisons between participants. To achieve this, he developed a Free Chain Association Task (two studies) and a Thought Process-Writing Argument Task (one study). In the Free Chain Association Task, participants were asked to write words associated with a stimulus word (e.g., "laboratory" or "building"). They could write more than one associated word per stimulus word. To test the effects of PO, all participants performed their sessions both in a group (3–5 participants per group) and alone (i.e., non-presence of others, or NPO); the session order was counterbalanced with a one-week interval between them. The results showed that 93% of participants produced more word associations in PO than when alone. Participants' associations were also divided into two different categories: Personal Associations (related to subjective experiences with the stimulus word) and Free-Rising Ideas (common associations). This categorization revealed that in NPO, participants produced more Personal Associations, while in PO they produced more Free-Rising Ideas. Allport (1920) interpreted these results as a signal that in NPO we are more subjective, while in PO we are more objective.

In the Thought Process-Writing Argument Task, Allport asked participants to write arguments about a given written passage; the results replicated those for the Association Task in that more arguments were produced in PO than in NPO. However, Allport noted that, despite the higher number of arguments produced in PO, the quality of those arguments was worse than those produced in NPO. This drew attention to a kind of trade-off between quantity and quality. Thus, Allport's (1920) study suggested that PO facilitates quantity but decreases quality.

Allport continued his approach to SF not only with the previously described tasks but also with the use of a Multiplication Task (Allport, 1924), which was essential for a better understanding of the effect. In this task, participants were presented with multiplications and their results, and were asked to decide whether the result was correct. Contrary to previous results, now participants in PO performed worse than NPO participants. Despite this evidence that social presence does not always

improve performance, but rather impairs it, Allport did not change the term Social Facilitation when naming the phenomena; thus, the term has been maintained, including the possibility of inhibition/impairment in PO.

The studies that Allport developed, and their results, introduced new ideas about the way PO and/or social contexts without competition can influence performance and behavior. From these studies emerged critical points that would become highly relevant in the future:

1) SF effects occur even without a competition aspect; this later introduced the question, "what is the minimal social condition required for SF effects to occur?" This helped the SF field to differentiate between SF and competition effects (Strube, 2005).

2) SF effects also occur with cognitive tasks, despite the bigger focus on SF in cognitive processes only taking place at the beginning of the 21st century.

3) PO promotes facilitation effects on quantity, but impairment effects on quality; this raised the question of whether the PO always is or is not a facilitator of performance, and which different tasks might be affected in different ways by PO.

4) In PO, there seems to exist a spreading out of associations, suggesting that the mind makes more associations in PO, but at the cost of losing original or subjective associations. This was the first suggestion that our minds work in a completely different way in PO versus NPO.

Later, Dashiell (1930) addressed Allport's conclusions about the quantity/quality trade-offs in social contexts, and the minimal conditions for inducing SF effects. He set up an experimental design in which each participant performed a set of tasks (e.g., Multiplication Task and Word Association) in different social contexts: a) alone, b) coaction, c) competition with another participant, and d) observed by two persons. In addition to these manipulations, Dashiell's design differed widely from Allport's studies in that he recorded the time required to complete the tasks. He concluded that the conditions of competition and observation promoted speed (i.e., less time for task completion), but when the task permitted evaluation of the accuracy, these two conditions had the worst performance compared with the alone and coaction conditions. Strangely, Dashiell did not observe any difference between the alone and coaction conditions. He attributed this result to a possible flaw in his experimental setting, which allowed participants who were doing the task alone to know that other participants were doing the same task elsewhere. This may have nullified the alone condition.

Dashiell's (1930) studies made three important contributions to the field:

1) The type of PO matters: observation and competition lead to faster and worse performances than those of coaction.

2) The SF effects are susceptible to a trade-off between speed and accuracy, where faster responses lead to less accuracy.

3) It is possible to make someone who is alone "feel" as if they are in PO; in other words, priming PO. This was later verified (e.g., Aiello & Douthitt, 2001; Hall & Henningsen, 2008; Anderson-Hanley, Snyder, Nimon, & Arciero, 2011; Uziel, 2016).

Despite Dashiell's and Allport's organized approaches to studying the effects of PO on participants' performances, researchers' conclusions in this early stage produced more questions than answers by generating highly ambiguous information about the main definition of SF effects and thus what the conditions are that define its occurrence.

The studies developed to this point created ambiguity about the direction of the PO effect: some studies reported positive effects on performance in PO (Facilitation), while other studies reported negative effects (Inhibition). Although the two directions were identified by Allport, he suggested that PO facilitates speed but decreases quality/accuracy (Allport, 1920, 1924; Dashiell, 1930). However, several studies reported positive and negative effects on performance using the same performance criteria (i.e., accuracy). For instance, Travis (1925) used a Hand-Eye Pursuit Coordination Test, in which participants coordinated their hand movements with their vision; he observed that in PO, participants made fewer errors than in NPO. Another example of positive effects of PO on performance can be found in the work of Bergum and Lehr (1963); participants observed a sequence of lights and pressed a button when they detected an error in the sequence. In PO, participants correctly detected more failures in the sequence than participants who did the task alone. By contrast, in studies using finger mazes (Husband, 1931; Pessin & Husband, 1933), which also require coordination, the better performance was observed in NPO—fewer errors in completing the maze. Pessin (1933) also observed that participants could memorize better nonsense syllables when alone. He concluded that the performance inhibition in PO was due to the interference created by distraction of being in PO, a term that later would be important for some SF theories. Ader and Tatum (1963) also presented results that pointed to the possibility that PO works as a distractor. In their study, participants received low-voltage shocks until they could find out how to stop it (a button); learning which button stopped the shocks happened faster in NPO than in PO.

The understanding that PO can promote either facilitation or inhibition led some authors to suggest that explanations of SF must account for the bi-directionality of the effect and others to defend the position that SF should be distinguished from a social inhibition effect (see, Guerin & Innes, 1984).

As noted by Guerin (1993), the majority of studies did not report the SF conditions correctly, which also contributed to an ambiguous definition of the SF effect. If an SF study wants to test the effect of PO on performance, it is critical to have at least one condition in which the participants' performance is measured in PO, and another condition in which the participants' performance is measured in NPO, completely alone. However, many studies include the presence of the experimenter in the so-called "alone" condition, a problem that, unfortunately, persists today. This lack of rigor could be one of the reasons for the ambiguous results. For instance, the presence of the experimenter in the room during the so-called "alone" condition could promote evaluation effects, leading to wrong conclusions about SF effects.

Results from different directions, different tasks, and different manipulations of PO, and a lack of rigor in the alone condition led to much confusion in the SF field at the time. Rosch (cited by Strauss, 2002) labeled this SF period as, "decades of experimental anarchy." SF effects were in need of theories, theories that would help generate predictions and also expose possible moderators (e.g., types of PO and/or types of tasks).

Social Facilitation Theories: The right moderator, the right mechanisms?

The need for a theory that could explain SF effects was addressed by Zajonc (1965). Zajonc was clearly aware of the conflicting evidence from various studies—some reporting better performance in PO and some reporting worse performance—and thus focused his theoretical approach on trying to account for it. In this process, he noticed that tasks that involved learning, such as learning nonsense syllables (Pessin, 1933), learning the right path to finish a maze (Husband, 1931; Pessin & Husband, 1933), and learning which button to push to avoid shocks (Ader & Tatum, 1963), resulted in impaired performance in PO. By contrast, tasks such as monitoring errors and coordination without the need to learn (Travis, 1925; Bergum & Lehr, 1963), led to better performance in PO. This observation led him to a theory that was inspired by the Hull-Spence Drive Theory (Spence, 1956).

The behaviorist Hull-Spence Drive Theory addressed the impact of motivation (drive) on animal performance. The theory postulates that when animals experience a specific motivation like fear or hunger, they are driven to emit a dominant response (i.e., a well-learned response). Zajonc

(1965) thought that PO was a possible motivational variable. If true, PO would cause a drive increase, which would arguably lead to a higher likelihood of dominant responses. However, because a well-learned response would be facilitated in a PO context, the unlearned responses (not dominant) would be inhibited, thus making it more difficult for individuals to provide a non-dominant response (i.e., not well-learned response). This account could explain why performance might be impaired in PO conditions when learning is experimentally induced (learning tasks), because the emission of a dominant response may interfere with the emission of the unlearned response. On the basis of this logic, Zajonc proposed classifying experimental tasks into two groups: simple tasks and complex tasks. Simple tasks would be those that demand dominant responses, and thus participants would perform better in PO. Complex tasks would be those that demand non-dominant responses, and thus participants would perform better in NPO. In sum, the type of task as a moderator of SF effects was a key aspect of Zajonc's theory, which would be adopted by future approaches.

Another key aspect of his theory is the idea that PO is a motivational variable, namely a source of general activation (arousal/drive). The assumption was that participants would be more aroused in PO. Zajonc (1965) did not present direct evidence that PO increases general activation in humans, rather, he stated that "the evidence which bears on the relationship between the PO and arousal is, unfortunately, only indirect" (p. 273). He maintained his hypothesis based on animal studies that show that increasing population density (i.e., number of animals in the same space) triggers the endocrine system (Thiessen, 1964; Lasagna & McCann, 1957; Mason & Brady, 1956). A question that then became relevant was, "what type of PO is necessary to trigger the expected level of arousal?" At the time, Zajonc (1965) only described two types of presence, coaction and audience (i.e., direct observation). Later, this opened a discussion in the field introduced by Cottrell's (1972) theory of Evaluation Apprehension, which is presented below. For Cottrell, the only way that PO could trigger arousal was if it created stress, that is, a type of PO that made participants feel evaluated. This led Zajonc (1980) to clarify his theory, arguing that the effect is not driven only by an expectation of an evaluation, but rather can be promoted by every type of PO. While the Cottrell's theory includes competition, audience, and evaluation, the effect, Zajonc asserted, can be promoted by the mere presence of another person. Thus, his main point became one of defending this position: the mere presence of another person is enough to create SF effects. This claim would distinguish Zajonc's Drive Theory from other arousal/drive-based theories.

The importance of Zajonc's theory was that, by postulating that simple and complex tasks work as moderators of SF effects, it could explain not only previous studies but also future studies on humans and animals (Wheeler & Davis, 1967; Tolman, 1967; Zajonc & Sales 1966, Zajonc, Heingartner & Herman, 1969; Zentall & Levine, 1972).

Zajonc's Drive Theory and its correlates established principles for future approaches:

1) SF occurs in the mere presence of other co-species—an assumption that is, however, challenged by alternative approaches.

2) SF can occur because of motivational factors—an assumption that is challenged by theories that offer a more cognitive explanation.

3) Type of task moderates SF effects, such that simple tasks (which rely on dominant responses) promote facilitation, and complex tasks (which rely on computation of new responses) promote inhibition.

However, the following points of Zajonc's theory still required clarification.

The first point of clarification was the definitions of arousal and drive, and the proof that they increase in PO. No study that had measured arousal levels in PO had linked those levels to participants' performance. This may have been because the constructs, arousal and drive, were not well-defined at the time. When presenting his theory, Zajonc (1965) used the terms interchangeably, without distinguishing one from the other. Additionally, he believed that stress and physiological activation (regardless of the system; e.g., cardiac or endocrine) were operationalizations of arousal. Only much later would this issue be approached theoretically and experimentally by Blascovich and collaborators (Blascovich, Mendes, Hunter, & Salomon, 1999), who viewed arousal as a general activation of different components of the physiologic system, which can assume different patterns of activation. Thus, contrary to Zajonc's view, Blascovich et al. (1999) approached arousal not as a "higher or lower" activation of a system, but rather a pattern of activation of different systems, which have different implications for the activation of a dominant response (which is discussed further in this chapter).

The second point that required clarification was the relative definitions of a simple and a complex task. Although type of task as a moderator could explain some effects, making some predictions possible, it is often difficult to define a task as simple or complex before seeing the results. The Stroop Task (Stroop, 1935), for instance, is simple in its demands, as it merely asks participants to indicate the color in which a word is printed (i.e., color naming). However, this naming process is

disrupted because the semantic meaning of the word is a different color (e.g., the word *red* printed in the color green). As Stroop (1935) concluded, reaction times are slower when participants try to name the ink color of words that have the opposite meaning of their printed color, because the reading process is more automatic (i.e., dominant) than the color-naming process. Thus, the activation of the automatic response should be operationalized as a dominant response. Furthermore, on the basis of Zajonc's theory, PO should cause participants to perform worse because reading is a well-learned response. However, currently, we know that the opposite happens (Huguet, Galvaing, Monteil, & Dumas, 1999; Klauer, Herdfordt, & Voss, 2008; Sharma, Booth, Brown, & Huguet, 2010): participants perform better in PO than in NPO. If the reading component of a Stroop Task is dominant, then the weight given to the reading process should be increased and the Stroop effect should increase in PO. Thus, the definition of tasks as simple or complex, and definitions of responses as well-learned and unlearned, are too simplistic to apply to all SF effects.

However, if Allport's experiments and his definition of SF effects were relevant to the field, Zajonc's Drive Theory was essential for providing a way to explain it. With Zajonc's theory, the SF field was no longer an orphan of a theory, which led other researchers to view it with increasing relevance.

The fruitfulness of Zajonc's approach was evident in the number of reactions it stimulated. One of the first reactions, already mentioned, was Cottrell's (1972) Evaluation Apprehension Theory. In accordance with Zajonc, Cottrell's Theory agrees with the effect of arousal/drive on dominant responses, and thus on task complexity as a moderator of the SF effect. His main goal was to go beyond Zajonc's approach and explain how arousal/drive increases in PO. He posited that it increases as a reaction, learned in past experiences (i.e., drive as a learned reaction to the PO), to the evaluative presence of others. His view was that participants who felt apprehensive in present evaluations had experienced a past negative evaluation in PO. Therefore, an increase in arousal/drive is triggered. Thus, Cottrell believed that for SF effects to occur, participants would have to be concerned about others' evaluations. This challenged Zajonc's (1965; 1980) assumption and results: that mere presence is enough to trigger SF effects. Support for the Evaluation Apprehension Theory was offered in a study by Cottrell and colleagues (Cottrell, Wack, Sekerak, & Rittle, 1968) that used a Pseudorecognition Task. The experiment manipulated participants' verbal habits by asking them to repeat nonsense words. The idea was to turn their response to some nonsense words into a dominant response by repeating them 10–25 times. Other nonsense words were maintained as non-dominant

by asking participants to repeat them only 1–5 times. Zajonc and Sales (1966) used the same task to show that participants in the presence of an audience (observers) did better at recognizing the words that were repeated more frequently than when they did the task without the audience present. This suggests an increase in well-learned (dominant) responses in the presence of an audience. However, Cottrell et al. (1968) aimed to show that the effect was caused by an audience evaluative effect and not a simple PO effect. To do this, they replicated the study but added an audience that could not observe the participants' performances. The authors contrasted participants' performances in three different conditions: alone, where the assigned participant did the task alone in a room; with an audience and thus the possibility of evaluation, where the assigned participant did the task in the presence of two spectators; and with a non-evaluating audience, where the assigned participant did the task in the presence of two blindfolded "spectators." Compared with the alone condition, the results revealed no effects of the audience when they did not observe the participant's behavior. However, participants evidenced better word recognition in the evaluation condition than in the alone condition. These results suggest, as Evaluation Apprehension postulates, that the SF effect only occurs in evaluative contexts, and that the mere presence of others is not enough to stimulate the effect.

Research by Cottrell et al. (1968) introduced inconsistency into the literature. It contradicted other studies that showed that the way participants performed when alone was not similar to the way they performed in PO, even when the PO condition was not evaluative. However, the null comparison between alone and PO conditions found by Cottrell et al. (1968) is not a universally replicable effect. For instance, Henchy and Glass (1968) replicated the Pseudorecognition Task by orthogonally manipulating presence and evaluation apprehension. Thus, the alone condition could be performed with and without participants' performances being recorded for further evaluation. In the audience condition, participants either performed the task in the presence of two experts or in the presence of two informal students from another university. By comparing each condition with the non-evaluative alone condition (NPO and No Evaluation), results showed that the presence of experts condition (PO and Evaluation) promoted higher SF effects, followed by the alone recorded condition (NPO and Evaluation); the student audience condition (PO and No Evaluation) showed the weakest SF effect, but was still present. The authors concluded that, although evaluation was an important factor, the mere presence of other students should not be ignored as a potential promoter of SF effects. That is, an effect, although weak, can also be produced by mere presence.

The null effect of the blind audience condition in Cottrell et al. (1968) was also not replicated by Rajeck, Ickes, Corcoran, and Lenerz (1977), who tested SF effects in maze tasks to find out whether a pure mere presence condition could produce a different performance than an alone condition. Rajeck et al. (1977) used similar conditions to Cottrell et al. (1968): NPO; audience/evaluation (i.e., a person who observed the participant's performance); and blindfolded audience/mere presence (i.e., a person who was present but could not observe the participant's performance). The results revealed more errors and faster reactions times in the audience/mere-presence condition than in the NPO condition, suggesting that mere presence is enough to cause SF effects. However, while participants assigned to audience/evaluation also had faster reaction times than those in NPO, the numbers of errors were equal between the two conditions. Thus, although these results clarify that SF effects in mere presence impact both speed and accuracy in maze tasks, they also show that mere-presence effects and evaluation effects may be two independent effects. As such, it is possible that different processes are at stake.

The definition of SF as a mere-presence effect was further supported by the results of a meta-analysis (241 studies) conducted by Bond and Titus (1983), which revealed that SF effects are unrelated to evaluation conditions. Thus, SF effects were defined, as Zajonc (1965; 1980) had argued, as presence effects, where mere presence is sufficient to trigger them. The meta-analysis made clear, moreover, that SF effects cannot be investigated by comparing mere presence with evaluation conditions, as Cottrell's work indicated (Cottrell et al., 1968; Cottrell, 1972). This reinforced the importance of a true alone condition (i.e., without the presence of the experimenter) in SF studies.

However, if mere presence promotes SF effects, it is also true that the impact of PO on performance occurs because of evaluation apprehension. Evaluation is not a necessary condition for SF effects, but it is sufficient to promote differences between PO and NPO conditions. Such differences should not be designated as SF, however, as they are not caused by PO itself (i.e., by just mere presence). The same is true for other effects that have PO as a necessary condition, such as social comparison (see Aiello & Douthitt, 2001). All social comparison theories share the assumption that PO causes participants to manage their behavior to give others a good impression. Aiello and Douthitt (2001) defined the following theories as examples of this phenomenon:

Self-Presentation Theory with drive (Baumeister, 1982) and without drive (Bond, 1982): these theories are based on Goffman's (1959) work, who postulated that individuals make an effort to manage the impressions they "give off" to please others. Baumeister (1982) extended this idea to

SF, proposing that PO would trigger individuals' drive to manage their public image, which would impact their task performances. He also pointed out the possibility that certain social conditions could trigger this drive more than others; the drive would be higher in evaluative conditions than in mere presence conditions. Despite having the same idea—that individuals strive to manage their impressions to please others— Bond (1982) did not use drive as a key factor for SF effects, but rather the concept of effort. In this way, he postulated that individuals in PO want to present an image of themselves as competent, which leads to more effort exerted during experimental tasks. Therefore, if the task is simple, their efforts would improve their performance; if complex, the embarrassment over some errors would lead to worse performance.

Objective Self-Awareness Theory (Duval & Wicklund, 1972): the basic assumption underlying this theory is that PO stimulates more awareness in participants of how others may see them. This leads to an increased focus on personal ideals and on efforts to avoid failing to achieve a goal. This produces a better performance in PO if the task is simple, but a worse performance if the task is complex.

Discrepancy-Reducing Feedback Loop (Carver & Scheier, 1981): this theory is based on Objective Self-Awareness theory, but is more detailed. Like the previous theory, it states that PO causes individuals to focus their attention inward, engaging a mechanism called a discrepancy-reducing feedback loop. This mechanism aims to minimize discrepancies between the individual's goal and what they are actually doing (i.e., my performance versus my ideal performance). For the individual to continue trying to reach an ideal performance, they must maintain a high expectation of achievement. This is easy with simple tasks where errors are rare; however, with complex tasks, the larger number of errors will lead the individual to think it is impossible to reach their goal, and thus tend to stop putting more effort into the task.

Because PO generates a context in which evaluation and social comparison are likely to emerge, researchers (e.g., Cottrell et al.,1968) have tended to see evaluation and social comparison as factors that explain SF effects. The position taken in this thesis, however, is that SF effects arise from mere presence (i.e., by PO itself); it is assumed, moreover, that this parsimonious view can bring clarity to the field. As such, Zajonc's Drive Theory (1965, 1980) directly addressed the SF phenomenon.

To summarize, several factors that co-occur with the presence of others, namely evaluation and social comparison, are likely to impact individuals' behaviors. However, these factors are not a

necessary condition for SF effects, and their effects should not be designated as SF, because they do not result from the mere presence of another person.

Despite the legitimate concerns introduced by evaluation and social comparison, one theory emerged to seriously challenge Zajonc's approach and is seen as its greatest rival: The Distraction-Conflict Theory and its derivatives (Sanders & Baron, 1975; Sanders, Baron, & Moore, 1978; Baron, Moore, & Sanders, 1978). The main claim of Distraction-Conflict Theory is that SF effects are promoted via distraction. In its first version (Sanders & Baron, 1975), the authors argued that by being distracted from their goal, individuals increase their motivation (drive/arousal) to attend to that goal, causing performance differences: specifically, facilitation in simple tasks and inhibition in complex tasks. By claiming the impact of PO over distractions, the theory offers a cognitive explanation of why drive increases in PO.

Sanders and Baron (1975) tested the hypothesis that distraction has "drive like" effects on task performance by manipulating distraction using procedures other than PO, and asking participants to perform a simple and a complex version of a "Copy Task." The simple version of the task consisted of copying (writing) a row of symbols. The complex version required participants to first translate the set being copied based on a key code (e.g., "1" is the code for "5") or to rewrite the symbols upside down and backwards (e.g. "6" would be recoded in "9"). The authors manipulated distraction by asking participants to listen for a signal that would be presented at the same time that they were performing the task. This made participants look away briefly from the task, shifting their attention between the "Copy Task" and listening to the signal. Their results indicated that participants in the distraction condition performed better (i.e., correct number of copied symbols) than those doing the simple task with no distraction. However, in the two studies presented in that paper (Sanders & Baron, 1975), this difference did not reach statistical significance for complex tasks, and the conclusion about whether distraction really could decrease performance in complex tasks was ambiguous. Despite this, these results are considered as support for the Distraction-Conflict Theory.

Later, Sanders et al. (1978) tested whether distraction could be caused by attentional shifts between the task and social stimuli. This approach assumed, again, that social comparison is a relevant mechanism for PO. Further, although they did not predict that social comparison was the only way to create distraction, these were the conditions in which they tested their hypothesis. They assumed that the relevant mechanism for causing SF would be the conflict created between attending to the task and to any irrelevant social stimuli. In the two experiments that they developed to test the

hypothesis, they used the Copy Task again, and orthogonally manipulated Social Condition (PO vs NPO) and the motivation to seek for social comparison. In the first experiment, the motivation to seek for social comparison was manipulated by telling participants either that the aim of the task was to measure their ability to defer gratification (high motivation for social comparison), or that it was to simply get their personal impressions of the task (low motivation for social comparison). In the second experiment, the manipulation was made by telling participants that others were or were not working on the same task—the expectation was that participants who believed they were working on the same tasks as others would be highly motivated to seek for social comparison, while those working on different tasks would be much less motivated to seek for social comparison. The results revealed that differences between NPO and PO conditions only occurred in the conditions with high motivation to seek social comparison. The authors argued that the drive only increases if there is the possibility of comparison; SF was once more understood as a social comparison effect, and not a mere presence effect in itself. However, if this was the case, what explanation is there for data (see Bond & Titus, 1983) showing that mere presence is enough to stimulate SF effects? In addition, contrary to Sanders and Baron's (1975) study, Sanders et al. (1978) did not find SF effects in complex tasks. This is a problematic result, given that the authors aimed to demonstrate that social comparison, and thus distraction, is a factor for SF effects in both simple and complex tasks.

Importantly, none of the foregoing studies confirmed the existence of a hypothetical level of distraction in PO. To address this gap, Baron, et al. (1978) ran another study in which they measured each participant's distraction levels after their performance in a Paired-Associate Learning Task (Spence, Farber, & McFann, 1956), performed either in an NPO or a PO (presence of an audience) condition. A simple and a complex version of the task was used: participants had to learn either a noncompetitional or a competitional list of paired word associates. A noncompetitional list of associative pairs maximizes the strength of association within pairs and, at same time, minimizes the association between members of different pairs (e.g., adept-skillful, barren-fruitless). A competitional list comprises associated pairs (e.g., barren-fruitless) and non-associated pairs (e.g., arid-grouchy, desert-leading), which, despite having no association between the members, are related to a member of another pair (having a between-members association; e.g., arid and desert). These features make this list more difficult to be learned well (i.e., complex task). Learning was performed in NPO for all participants, afterwards half of the participants were tested in an audience condition, while the other half were tested in NPO. Results confirmed the standard SF effects: better performance in PO than in NPO for the noncompetitional list (simple task), and worse performance in PO than in NPO for

the competitional list (complex task). Levels of distraction were measured in different ways: by directly asking participants how frequently they focused their attention on the task; by asking participants to recall different features of the associative pairs (e.g., first letter was lower case or printed in red); and, using the number of errors committed as an index of distraction. All participants in PO evidenced higher distraction levels than participants in NPO. This result suggests that distraction is a factor that can improve or impair task performance, depending on task complexity, assuming that the data are interpreted as evidence that participants' performance was promoted by an arousal caused by distraction.

Later on, Sanders (1981) carefully analyzed how results obtained in the field helped to support different SF explanations, contrasting Zajonc's (1965, 1980) theory and the Distraction-Conflict Theory (Sanders et al., 1978). He concluded that no model by itself completely explained all the reported results, and proposed a synthesis model: the Attentional Process Model. This model assumes that SF effects can occur via three different pathways that are not mutually exclusive: reflexive response (activation of a dominant response), learned anticipation (learned drive) of positive or negative outcomes, or/and distraction-conflict. Although the integration of all these assumptions allowed for a better account of the available results, Sanders failed to explain when and how the processes that defined each pathway occurred (see Guerin & Innes, 1984).

One caveat of the Attentional Process Model, and the models that support it, regards the relationship between attention and arousal, that is, the assumption that distraction impacts arousal. No direct evidence is offered for this assumption. Additionally, an alternative view was offered by Easterbrook (1959), who showed that the direction can be exactly the opposite (i.e., arousal affects distraction). Evidence for this can be found, for instance, in Zaffy and Bruning's (1966) study, which asked participants to learn an implicit rule underlying a correct response in a condition where relevant, irrelevant, or no cues were offered. Results showed that participants with more arousal/higher drive, measured by an anxiety scale, performed better when cues were irrelevant for the task. However, in conditions where a relevant cue was offered, the participants with more anxiety had the worst performance. This indicates that arousal led to the use of fewer cues; thus, when the cues are irrelevant, arousal is a facilitator of performance, because participants are not attending to the irrelevant/disturbing information. When the cues are relevant, high-arousal participants do not benefit, but rather perform worse in comparison with low-arousal participants, who use and benefit from the presence of relevant cues.

Later, Bruning, Capage, Kozuh, Young, and Young (1968) replicated Zaffy and Bruning's (1966) study using different social contexts. In one experiment, the researcher peered over the participant's shoulder to record the response (High-drive condition), or sat behind a screen (Low-drive condition). In another experiment, the researcher, who in this case remained behind the screen, indicated to the participant that their performance was: below the average (High-drive condition) or very satisfactory (Low-drive condition). The results showed, again, that participants in the high-drive condition performed better than participants in the low-drive condition when the cues were irrelevant to the task; the opposite pattern was observed when the cues were relevant to the task: better performance for participants in the low-drive condition. However, if this work is anchored in Easterbook's (1959) model, which relates arousal to attention, all that Bruning et al. (1968) did was add one more possible moderator to the Drive/Arousal Theory (Zajonc, 1965, 1980): type of task cues (i.e., relevant or irrelevant to the task objective).

These empirical approaches, which have further clarified how the Easterbrook Model (1959) helps explain SF effects, ended up challenging arousal as the mediator through which distraction promotes SF effects. In addition, its methods and data demanded a review of the explanation for how the SF effect occurred, since PO was defined in many different ways: as a source of arousal, as a distraction, and as a context that modulates the use of relevant and irrelevant cues.

Studies using a cognitive approach ended up challenging arousal as the mediator through which distraction promotes SF effects. The focus on the Easterbrook Model (1959) and its assumptions about perceptual overload and attentional priority relating to a narrowing mechanism of the range of the perceived stimuli (Cohen, 1978) challenged Baron (1986) to leave out assumptions regarding arousal and to reformulate the Distraction-Conflict Theory. Baron (1986) assumed that distraction–conflict observed in PO imposed a cognitive overload on participants that may also have produced attentional focusing. Attentional focusing is defined as a narrowing of attention that causes participants to attend better to what is central and to neglect what is not. In this second version of the Distraction-Conflict Theory, also called the Overload Hypothesis (Strauss, 2002) in the literature, there is no role for arousal in explaining the SF effects. Baron (1986) proposed that the PO is a distraction context, and hence creates a conflict between the attention to the social context and the level of attention needed to perform the task. This conflict creates an overload that restricts individuals' cognitive focus. This focus, or narrowed attention, leads them to attend more to cues that are relevant/central to the task at hand, but at the expense of not attending to the peripheral cues.

This assumption—that PO promotes a narrowing of attention—matches the mechanism presented in Easterbrook's Model (1959). However, Baron (1986) also used Cohen's (1978) ideas (Cohen's approach was more purely cognitive), and erased the drive/arousal from the model as either antecedent or consequent, assuming that the cognitive overload itself led to narrowing attention. Additionally, and also in contrast to the Easterbrook Model, Baron specified that performance is impacted by PO because it draws on the individual's attentional resources and thereby narrows their attention.

Baron's understanding of the role that task complexity plays in SF effects rested on the fact that complex tasks have numerous relevant features; because they have more relevant features, complex tasks consume cognitive resources, and thus increase the cognitive load that is already imposed by PO. Because of the cognitive load, participants cannot attend to all the stimuli, which narrows their attention; thus, they do not attend to all relevant stimuli and so their performance decreases. By contrast, simple tasks have few relevant stimuli and thus the narrowed attention created by PO helps participants ignore irrelevant stimuli, which allows a better performance.

However, although Baron (1986) suggested that PO narrows attention because of cognitive overload, he did not clarify why PO creates such overload. The explanation left open by his approach was one that went back to the same factors previously identified as arising from PO and repeatedly claimed to explain the SF effects (i.e., because individuals engage in social comparison, evaluations, etc.). Another caveat of the Overload Hypothesis was its capacity to predict various types of PO effects; that is, the theory seemed less likely to predict some types of social presence (e.g., coaction or mere presence) than others that, arguably, are more disruptive (e.g., evaluation or competition).

The difficulty in developing a simple theory to explain the SF effects suggests that there are several factors associated with the presence of others in a given context that can exert an effect on an individual's performance. Indeed, investigating the effects of PO are more complex than initially thought, and as scientific phenomena, they cannot be explained with a simple and direct theory. Thus, it is important to both clarify misunderstandings of the theoretical concepts previously used and explore additional, potentially relevant, factors for explaining such effects.

Old problems, new problems: The path for clarification.

From this brief overview of SF studies, a set of old problems that need clarification can be identified. Below is a review of some of these old problems, which are related to the lack of

identification of possible moderators of SF, and are related to problematic confounds of effects related to PO in the current literature.

One of the misunderstandings around SF models is their use of an arousal/drive concept as a core mechanism of SF effects; in fact, there is no clear definition of this concept. The idea of arousal as an explanatory factor for SF effects started with the idea that PO leads to more physiological activation, possibly in the endocrine system (Zajonc, 1965). However, this system hypothesis was subsequently forgotten. The remaining approaches referred solely to a "generalized arousal," viewed as general increased activity in all physiological systems (e.g., cardiovascular system and respiratory system). This, however, was not supported by the Bond and Titus (1983) meta-analysis, which showed that the only physiological measure that is affected in SF by PO is palm sweating, and only during complex tasks. These results clearly challenge the idea of arousal as the core mechanism that leads to SF effects. The concept of "generalized arousal," moreover, is strongly challenged when considering that most arousal manipulations cause some physiological measures to increase while making others decrease (Sanders, 1981; Carver & Scheier, 1981; Cacioppo & Tassinary, 1990). This led Blascovich et al. (1999) to apply the Biopsychosocial Model of Challenge and Threat (Tomaka, Blascovich, Kibler, & Ernst, 1997) to SF effects.

This Model postulates that challenge and threat states are activated in goal-relevant situations (i.e., situations where self-evaluations are seen as important). When individuals believe that their performance is important and simultaneously think that they have the resources to fit the task demands (i.e., capacity to respond correctly), they experience a challenge state. However, a threat state occurs when participants believe that their performance is important but do not feel they have enough resources to fit the task demands. If PO impacts arousal, it should map onto these states either by increasing both or by selectively promoting one over the other. This is a directly testable hypothesis as each state is followed by a different physiological pattern. While both challenge and threat states are characterized by increased cardiac responses, the former involves a simultaneous decrease in vascular resistance, while the latter involves a simultaneous increase. Thus, the concept of "generalized arousal" should not be defined as a continuous concept of physiological activation that increases or decreases; rather, it should be defined as an activation pattern in different systems (i.e., cardiac and vascular resistance), which can translate in different categorical "types of arousal" (i.e., challenge and threat).

Blascovich et al. (1999) addressed various impacts of arousal with the support of SF effects that are moderated by task complexity. In other words, they predicted different effects for simple (well-learned) versus complex (unlearned) tasks. Following their rationale, PO causes participants to feel the task as more goal-relevant, which leads to a "challenge pattern" in well-learned tasks, and to a "threat pattern" in unlearned tasks. They tested this hypothesis by manipulating the SF conditions (PO, with two people observing vs NPO, where participants were alone) and the task mastering. All participants learned to master either a Number-Categorization Task or a Pattern-Learning Task through a practice block. In the experimental conditions, participants either performed the task they had mastered or a new task. The researchers measured both performance and physiological patterns and compared well-learned tasks with unlearned tasks. Only participants in PO (vs. NPO) showed a physiological activation pattern changed from their baselines; however, this activation showed a challenge pattern for those performing a well-learned task, but a threat pattern for those performing an unlearned task.

These data help clarify results that were previously interpreted as supporting Zajonc's (1965) assumptions that PO leads to more "generalized arousal" and thus to more dominant/well-learned responses. By contrast, the Biopsychosocial Model of Challenge and Threat suggests that PO leads to more goal-relevance, which, in turn, induces a cognitive appraisal between individuals' resources and task demands; this can result in a challenge or threat evaluation that activates the respective physiological patterns.

The idea that the impact of PO on physiological activation is dependent upon an individual's cognitive appraisal opened the discussion to the possible relevance of other individual differences in SF effects. Traditionally, neither Social Psychology nor Social Cognition had looked for moderation effects by individual factors, and, thus, little attention had been given to the impact of individual differences in SF effects, despite supporting evidence offered by earlier studies. For instance, Triplett (1898) and Allport (1920) reported that some of their subjects (around 7%) had opposite results (i.e., better performance in NPO than in PO), but no further focus was given to that phenomenon (see Uziel, 2007). However, individual differences as moderators of how people approach social situations (and thus, in PO), have been discussed in some theories of personality. One example is the Sociometer Theory (Leary & Baumeister, 2000), which supposes that individual's self-esteem level guides their approach to a social context. Additionally, Grant and Dajee (2003) suggested that introverts are more affected by SF effects than extroverts when doing multiplication tasks. Uziel (2007) assumed,

therefore, that specific personality traits predispose individuals to a positive or negative orientation towards social contexts, which would moderate PO effects. To test his assumption, he conducted a meta-analysis with SF studies that measured self-esteem, extroversion, and/or neuroticism. Results revealed that participants with a positive orientation toward PO (i.e., high self-esteem and/or high levels of extroversion) performed better in PO than in NPO. Conversely, participants with a negative orientation toward PO (i.e., low self-esteem and/or high levels of neuroticism) performed worse in PO than in NPO. In addition, no impact of task complexity on these effects was observed.

Together, Uziel (2007) and Blascovich et al. (1999) showed that the appraisal of the context and the task are important moderators for SF effects. From Uziel (2007) we know that participants' appraisals of social contexts as positive or negative moderates the SF effect. From Blascovich et at., (1999) we know that how participants feel about the task—whether they have the necessary resources (i.e., is it a challenge or a threat?)—determines the presence of a facilitation or an inhibition effect. The problem of relying on this conclusion, however, lies in the confounds typically made in PO studies, which adds to the methodological chaos often found in the SF literature. Uziel's (2007) meta-analysis of 20 studies only included two with a real NPO condition. That is, only two of the analyzed studies had a condition in which the participants performed the task alone (without the presence of the experimenter) and also a condition in which they did the task in PO (even then, there was a mix of mere presence and co-action conditions with other social contexts). After Zajonc's (1965, 1980) advertence to the need to incorporate a real alone condition in study designs, authors continued to ignore the importance of the social isolation condition, hence confounding other social effects with those that can only be connected to SF effects.

The conceptualization of SF as a mere presence effect is another problematic issue across different approaches in the field. While PO is a necessary condition for the occurrence of other social presence effects, many different variables may covary with social presence and exert their own effects. Importantly, these variables should not be confounded with SF, as such effects can either promote new effects that could then be confounded with SF or can qualify SF in different ways. One example is the Conformity effect in social groups, which occurs when people construct or follow norms that they observe in other people's behaviors with the aim of determining the correct behavior for themselves (Ash 1951). A famous conformity study, conducted by Sherif in 1936, is the Autokinetic Illusion. This illusion consists in the sensation that a light is moving around in a dark room; the truth is, without a reference point to fix our eyes in the darkness, our eye movements create the impression

of a moving light. Sherif asked participants to estimate how many inches the light moved. Half made their estimates in a group setting across many trials, while the other half made theirs alone. The results revealed that in a group, participants were prone to give similar estimates across all the trials, as if they were collectively building a common rule/norm to estimate the distance.

While conformity effects do occur in social contexts, they are not SF effects. They are effects that are dependent upon the nature of the relationships built between different members of a group, and that focus on individual behavior relative to group behavior. By contrast, the participants in an SF study never listen to other participants' responses and also perform their tasks in isolation. However, in the control condition of conformity studies, participants offer their responses directly to the experimenters and are thus engaged in social interaction with them. In other words, in a conformity study, there is no pure alone condition. Although the distinction is clear, the literature has continued to confound the two effects; for instance, in Vaughan and Hogg's (2005) conformity chapter, SF studies are cited as representing "a type of conformity."

SF is also confounded with Accountability effects, which address public vs private social contexts; it demonstrates that people have different attitudes if they know in advance that their responses will be public or private. This effect is illustrated, for instance, in Maass and Clark's (1983, 1986) studies in which participants demonstrated different attitudes towards gay rights depending on whether they perceived their responses as public or private. Explanations of this effect (e.g., Moscovic & Personnaz, 1980; 1986) indicate that in a public context, participants' responses tend to conform more to views that reflect norms of the social majority than they do when answering in private. This implies some type of anticipation of interaction that should not be a feature in an SF study. Furthermore, accountability studies do not even require the physical presence of others. Despite this, some authors confound the two effects and use accountability contexts (Public and Private) as SF Conditions (PO and NPO; e.g., Lambert et al., 2003). Another issue in confounding results of accountability with SF is that the NPO and private conditions used in the two types of experiments typically do not match. In most accountability studies, participants in the private conditions are in the presence of others (at least that of the experimenter) but receive instructions that imply that their responses are anonymous. Thus, there is no alone condition.

SF is also commonly confounded with the Social Loafing effect. Social Loafing describes the decrease in a participant's performance during collective tasks compared with their performance during individual tasks. For instance, using a Rope-Pull Task, Ringelmann (1913) showed that the

force exerted by each person decreased as a function of the group size (i.e., one, two, three, or eight). In other words, more people lead to less effort contributed to a collective task by each individual. Although this context clearly differentiates it from an SF context, in which the individual's task is not performed collectively, authors tend to confound them, claiming that SF effects occur in Social Loafing contexts such that participants perform worse in PO, which illustrates a social inhibition effect (e.g., Harkins, 1987). The confounding of the two effects is exacerbated by the fact that the detection of Social Loafing does not require an alone condition; furthermore, the "control conditions" in these studies typically consist of participants performing the task "alone" while being watched/evaluated by the experimenter.

The final example of an effect that is often confounded with SF is the Group Polarization effect (Myers & Lamm, 1976; Isenberg, 1986), which shows that people tend to be more extreme in their position, choices, or decisions, after a group discussion. Studies of this phenomenon have included political decisions (Janis, 1991), investment decisions (Whyte, 1993), and jury decisions (Bray & Noble, 1978). Again, this effect is dependent upon participants interacting with others. In Group Polarization studies, a participant's position is measured before and after a group discussion. Thus, this experimental paradigm does not share the most important features of the SF experimental paradigm: the manipulation of PO. Rather, participants are always in a group context and they encourage group discussions, which never occurs in the contexts in which SF effects are studied.

These claims are hardly the first ones to be made regarding the problems with the foregoing confounds. Authors such as Markus (1978), Bond and Titus (1983), and Guerin (1993), have criticized the field, pointing out that many SF studies confound SF effects with other social context effects because of their lack of a proper experimental control. Additionally, by including the physical presence of the experimenter, many studies fail to meet the proper definition of an alone experimental condition. This is problematic not only because it represents a "mere presence" condition but also because it is likely creating an evaluative context.

To isolate the SF effect from other effects, researchers must set up a PO condition in which no interaction between participants occurs either before or during the experiment, and must set up an NPO condition in which participants are completely alone in the room (without the presence of the experimenter). Furthermore, because the effect relies on an awareness of others' presence, data should not be collected in labs (settings) in which isolated cubicles are assigned to each participant; while some participants may feel isolated from others, some can be distinctly aware of the presence of others

in other cubicles. It has already been shown that in isolation it is easy to prime PO through virtual human faces (Park & Catrambone, 2007), interactive animations (Hall & Henningsen, 2008), eye images (Na & Kitayama, 2012), and through imagining social situations (Puntoni & Tavassoli, 2007; Figueira & Garcia-Marques, 2019). Differences in these experimental procedures are likely contributing to the diversity of data and effects that are found under the umbrella of SF manipulations.

Thus, to maintain rigor, SF effects should be isolated from other effects that can occur in a social context. To achieve this, studies must compare a pure PO condition (i.e., no interaction between participants) with a pure NPO condition (i.e., participants use the same room but completely alone; explicitly without the presence of the experimenter).

The diversity of experimental procedures addressing SF effects and the range of possible confounds with other effects leads to the question: how should we interpret previous research on SF? This is a challenge, especially when the details of the experimental environment are not fully described (e.g., use of cubicles or open rooms, coaction or alone, and with or without the presence of the experimenter). Thus, even after a century of experiments we do not know what the real boundary conditions of the effect are, or its antecedents and moderators.

Another difficulty that arises in reading the conjoint information provided in SF studies reported in the literature is their circumscribed focus on performance. Most of the studies focus simply on task performance levels, and thereby stick to Allport's (1954) original aim of studying "what change in an individual's normal solitary performance occurs when other people are present" (p. 46). However, the concept of performance itself is relative (Aielo & Douthitt, 2001). The SF effect arose in the context of quantified performance (e.g., speed, time, or number of associates) before Allport introduced the concept of performance quality (e.g., quality, creativity, and subjective responses; Allport, 1920, 1924). Landers and McCullagh (1976) have asserted, moreover, that Zajonc's theory (1965) only works for performance quantification, but not for performance quality. This would suggest that the SF effect is merely a matter of performance quantification, and is not related to how the performance is being produced.

To better understand the immediate antecedents to SF effects on performance, the field must aim for a more open perspective that goes beyond performance (see Guerin & Innes, 1984; Thomas et al., 2002). This point was partially made after Zajonc's theory (1965, 1980) had introduced the arousal/drive mechanism to the SF field as an explanation of SF; however, few studies addressed the impact of PO on arousal. Additionally, the emergence of cognitive explanations (Bruning et al., 1968;

Sanders & Baron, 1975; Sanders et al., 1978; Baron et al., 1978; Baron, 1986) may have implied a change of focus from performance to the internal processes that underly it. Only more recently, however, have scholars begun to follow up on this approach. A new generation of researchers (1999–2019) have been addressing the specific cognitive mechanisms that sustain the SF effects. They are reviewed below.

Attention, Distraction, and Executive Control, plus Context-Sensitivity effects: A new SF era.

The focus in SF research recently changed from behavioral effects to mind mechanisms; from qualifying performance to understanding the cognitive mechanisms that promote such effects.

It was the Overload Hypothesis (Baron, 1986) that triggered the study of cognitive effects on SF. Huguet et al. (1999) approached the hypothesis by directly comparing it with other attention-narrowing models (Easterbrook, 1959; Cohen, 1978) and tested for the first time the assumption that PO causes a narrowed focus of attention. In addition, this hypothesis was directly contrasted with the Zajonc view (1965, 1980), which assumes that PO increases the use of dominant responses within a Stroop task (Stroop, 1935). Following up on the idea that PO narrows attention depending on the task, other researchers (Wagstaff et al., 2008; Yu & Wu, 2015; Belletier et al., 2015; Hobson & Inzlicht, 2016) tested SF effects using either attentional or distraction mechanisms, to determine how PO interferes with general executive functions of the working memory system (namely, focusing and sustaining attention, updating, inhibition, encoding, and retrieval; Baddeley & Hitch, 1974). Importantly, Fonseca and Garcia-Marques (2013) opened an additional door along this SF avenue by suggesting that PO may also impact participants' sensitivity to the context such that in a social environment, their responses take the context more into account, thus spreading their attention.

The number of studies associated with this new approach that are being developed remains small, which allows for detailed descriptions (below) of those conducted with human subjects. At the end of the section, the relationships among the studies are discussed, as well as the possible integration of their results into one new theoretical framework.

Huguet, Galvaing, Monteil, and Dumas (1999)

Huguet et al. (1999) showed that in PO, compared with NPO, participants could more easily ignore irrelevant stimuli in a Stroop task. The authors believed this indicated that in PO, their participants could more easily narrow their attention to the relevant stimuli, which led to less Stroop interference.

In the first of two studies, social conditions were manipulated between participants in four different conditions: 1) NPO - participants did the task alone; 2) Inattentive-Busy PO - participants did the task individually with the presence of a confederate who was reading a book; 3) Invisible PO - the confederate sat behind the participant; 4) Attentive Audience - the majority of the time, the confederate was looking at the participant's computer screen. In each of these conditions, participants performed a Stroop Task in which different incongruent and neutral trials were randomly presented. In incongruent trials, both color names (classical Stroop stimuli: *red* written in blue) and semantic associates of a color were presented in an incongruent color (e.g., *blood* written in yellow). In neutral trials, non-semantic stimuli were presented in a color (e.g. ++++ written in blue), with the goal of reducing evaluative effects (which could be an alternative explanation for the effect). The authors told participants that the study's aim was to elicit their general impression of the task. As such, the computer was supposedly not recording their performance. At the end of the Stroop-like task, participants performed a recognition test of the words presented to determine whether the automatic interference of reading words was less in the PO conditions. Additionally, participants reported their levels of perceived distraction to assess whether PO increased cognitive overload. Results revealed less Stroop interference in two of the PO conditions than in the NPO condition. No difference was found between the NPO condition and the Inattentive-Busy PO condition. Additionally, levels of Stroop interference were related to levels of correct recognition, suggesting that the activation of the word meaning during the Stroop task left a memory trace. Furthermore, recognition levels were lowest in PO, which suggested less attention to the word meaning, and so a possible narrowing attention mechanism. Despite the lower level of recognition in PO, no effects occurred for the distraction measure.

Besides supporting the Overload Hypothesis, these results discarded the predictions that had been based on Drive Theory (Zajonc, 1965, 1980). Drive Theory predicted that in any task that participants perform in PO, they are more likely to offer a dominant response. In a Stroop task, the dominant response is the process of reading the word (e.g., automatically identifying the word meaning). Thus, according to the Drive/Arousal Theory, participants in PO should experience more interference. In fact, the opposite occurred; as the Overload Hypothesis (Baron, 1986) predicts, participants in PO evidence less interference during a Stroop task. While a Stroop Task can be seen as a measure of narrowing attention (Agnew & Agnew, 1963) because it focuses participants' attention on the relevant stimuli (i.e., the word color), the literature also defines a Stroop task as a measure of

executive control function (e.g., Ozonoff & Jensen, 1999; Phillips, Bull, Adams, & Fraser, 2002), a relevant mechanism for studies presented below (Augustinova & Ferrand, 2012; Wagstaff et al., 2008).

In the second study, Huguet et al. (1999) tried to replicate their effect by contrasting NPO with three coaction contexts: 1) with a slower coactor; 2) with a similar coactor; and, 3) with a faster coactor. Results were similar: more interference in the NPO condition than in the coaction conditions with a similar or faster coactor. No differences were found between the NPO condition and the slower coactor condition. Again, the results supported the Overload Hypothesis (Baron, 1986), wherein PO leads to the allocation of attention to central cues (word color) and the neglect of peripheral cues (word meaning), but only when PO is unpredictable or when the coactor is similar or faster. Because there was no SF in the slower coactor condition, this study suggests that mere presence is not enough to promote SF effects.

Klauer, Herfordt, and Voss (2008)

This paper addressed the possibility that instructions given by Huguet et al. (1999) to their participants promoted their results artificially. Performance of a Stroop task with a goal of forming a "general impression" seemed to produce higher levels of Stroop interference in the NPO condition (around 170 ms) than those usually reported in the literature. Klauer, Herfordt, and Voss (2008) wondered whether the impression formation instructions may have produced longer inspection times for the incongruent stimuli than for the neutral stimuli, specifically in PO. By replicating the studies (i.e., similar materials used in Huguet et al., 1999), but also using classic Stroop instructions (i.e., without the "general impression" instruction), these authors showed that under the "general impression" instructions setting (used by Huguet et al., 1999), participants in NPO tended to spend more time on incongruent words (e.g., *sky* written in red) but were faster with neutral stimuli (e.g., ++++). The supposition for this was that neutral stimuli have no semantic qualities to be analyzed and thus gave participants no reason to form an impression. However, when the neutral stimuli were also words (e.g., *table* written in red) the time spent was equal to that spent on the incongruent words. This suggests that in NPO, when participants are faced with the two demands of "general impression" and the Stroop Task itself, they privilege the former. In PO, however, participants focused on the Stroop task itself and not on the "general impressions" of the task. The authors' conclusion was that the reduction in interference scores in PO was not a function of attentional focusing on the task trials' features, which would have made it easier for participants to screen out the distracting word meaning, but of the specific experimental conditions created by Huguet et al. (1999). These conditions artificially

promoted more Stroop interference in NPO by making participants more introspective in trials where the targets were words. The authors further suggested that PO increases monitoring through task selection; participants were better at focusing their attention and hence at prioritizing the focused task (i.e., Stroop task) in PO, while in NPO, individual attention may be divided between the Stroop task itself and the impression formation of target word trials. Even so, a careful reading of the results of Klauer et al. (2008) reveals a marginal effect between PO and NPO in the same direction, as predicted by Huguet et al. (1999), when all the specific experimental conditions in their studies are controlled for. This did not remove the possibility of SF effects on a Stroop task through narrowing attention or executive control mechanisms.

Sharma, Booth, Brown, and Huguet (2010)

Assuming that PO affects Stroop interference, these authors tested whether this occurred because PO leads to a general narrowing of attention associated with an early selection mechanism, or whether PO causes an increase in monitoring, and hence an inhibition mechanism that encompasses the detection of the inference (between the word color and its meaning) and acts against it. Because this inhibition is a control process that involves both detection and inhibition of the word meaning, reaction times (RTs) should be slower for detection, compared with the (faster) RTs associated with detection in an early selection mechanism.

In this experiment, participants performed a Stroop task with classic incongruent trials (e.g., *red* displayed in blue and *blue* displayed in yellow) and neutral trials (e.g., ++++ displayed in blue); the within-response stimulus interval (RSI) was also manipulated to control the conditions that disfavor (RSI = 32 ms) and favor (RSI = 1000 ms) later control mechanisms. SF for between-participants conditions was created by having participants perform the task alone (NPO) or with a confederate present (PO). Results replicated previous data (i.e., PO showed less interference than NPO) but only for long RSIs. In addition, by using Cumulative Distribution Frequency (CDF; a method to split RT distributions into bins from faster to slower RTs; see Ratcliff, 1979; Houghton & Grange, 2011), the authors showed that the difference between NPO and PO only occurred at the slower RTs. The authors explained that whereas participants in NPO experience a pattern of increased interference throughout the RTs (i.e., less interference for faster responses but higher interference for slower responses), participants in PO evidence a decrease in Stroop interference for slower RTs. These data suggest that the mechanism that is reducing Stroop interference in PO is not an attentional one associated with a greater capacity to neglect the word meaning through narrowing attention (i.e., ignore

a peripheral cue through an early selection process), but a later and reactive process of inhibition/correction, which requires time and is more closely related to an executive control function. However, this interpretation does not align with the observation that participants in NPO are always slower in responding and that no differences in number of errors are found between the two conditions.

Augustinova and Ferrand (2012)

Augustinova and Ferrand (2012) extended and clarified previous results (Huguet et al., 1999; Klauer et al., 2008; Sharma et al., 2010) by addressing the hypothesis that the interference in a Stroop task may arise from two independent sources: semantic interference caused by the incompatibility between the color of the word and its meaning (e.g., *blue* written in red), and response competition between the word meaning and the keyboard (e.g., if the target word is *blue* displayed in red, the correct key to press is blue). This response competition assumes that effort is required to ignore the red key when giving the correct response (see Neely & Kahan, 2001; De Houwer, 2003; Schmidt & Cheesman, 2005). Thus, SF effects on a Stroop task may occur through both or only one of these specific pathways.

In two experiments, the authors used the same stimuli as those in previous studies: standard incongruent words (e.g., *blue* displayed in red), and semantic incongruent words (e.g., *sky* displayed in green). However, instead of integrating them into one task, they separated the stimuli into two different tasks. This was because semantic incongruent words are free from response competition (see Neely & Kahan, 2001), so using semantic and standard incongruent stimuli on different participants made it possible to isolate the semantic component of the Stroop interference from the response competition (see Neely & Kahan, 2001). Results revealed that PO only reduced the Stroop interference in task versions that used standard incongruent words. The authors assumed that the mechanism beyond the SF on Stroop interference was one of response competition; they concluded that this data refuted the narrowing attention mechanism as an explanation for SF effects in Stroop tasks. They claimed, rather, that SF effects arise from the use of a control mechanism that facilitates the response competition resolution, an interpretation that is congruent with executive control function.

Muller, Atzeni, and Butera (2004)

This paper addressed the narrowing attention mechanism by studying the effects of mere presence (and also of social comparison) on the illusory conjunction effect. The illusory conjunction

effect describes the tendency to imagine or see a target object when, in fact, only primitive characteristics of the object are presented (Treisman, 1988; e.g., on a computer screen, a participant identifies the target $ inside the primitive characteristics, "l" and "S", when in fact the target is absent). The effect is supposedly caused by a lack of attentional processing of the primitive visual aspects of the target. Social conditions were manipulated using mere presence (i.e., the presence of a confederate as co-actor) and NPO conditions (i.e., participants were alone in the cubicle during all experimental phases). Results of experiment 1 revealed that in NPO, participants made more conjunctive errors than in PO, which suggested that PO, via distraction, narrowed their attention to relevant stimuli (as suggested by Baron, 1986). Although the conditions of social comparison used in experiment 2 were not SF as defined in this thesis, they facilitated the understanding of the narrowed attention mechanism, which occurs in PO through distraction. Results of this study illustrate that Upward Comparison instructions (e.g., "You have made more errors than your colleague"), which are more disruptive/distracting (Sanders et al., 1978), reduces the number of conjunctive errors compared with Downward Comparison instructions (e.g., "You have made fewer errors than your colleague"). With respect to the goal of this thesis, the results of Muller et al. (2004) suggest that PO promotes narrowing attention, making individuals less prone to illusory conjunction effects that rely on undesirable interference.

Muller and Butera (2007)

In this study, comprising five experiments, the authors offered a replication of the impact of SF on illusory conjunctions and further explored conditions such as social comparison and evaluation of threat (see Cottrell, 1972; Evaluation Apprehension Theory). In experiment 1 and 2, the authors manipulated SF conditions (i.e., isolation vs co-action with one confederate), and thus replicated their previous work; results indicated fewer conjunctive errors in PO (no differences occurred when analysis was performed on the non-conjunctive error rate). Their results for social comparison and threat evaluation also indicate that these factors promote, via narrowing attention, fewer conjunctive errors.

Wühr and Huestegge (2010)

This study tested whether the presence of the experimenter could reduce a cueing effect through narrowing attention. Participants performed a Spatial-Cueing task, either always in the presence of the experimenter (mere presence condition) or split into two parts: first in the presence of the experimenter and then alone. The Spatial-Cueing Task (Posner, 1978, 1980; Posner & Cohen, 1984) consists of identifying the position of a target (e.g., letter "H") at four possible positions on the

screen (left, right, up, or down) while the other positions are filled by foils (e.g., "T", "M", or "L"). Before the target and foils are presented, an arrow appears in the center of the screen pointing to one position. For half of the trials, the arrow is a central cue and is thus cueing the target, while for the other half, the arrow is a peripheral cue, pointing to a potential location of the target. Faster RTs result when the cue is central, but longer RTs are produced when the cue is not central because the attention must be shifted to the target position—a phenomenon referred to as the cueing effect. Results showed that PO reduced cueing effects, possibly because it narrowed participants' attention. Experiment 2 tested whether this narrowing attention mechanism in PO depended on individuals' cognitive resources. Thus, they exchanged the arrow cue for a simple line, which did not have any symbolic directional value and hence no predictive value for the subsequent target location (i.e., arrows point but lines do not). According to the literature, this change creates a condition in which no driver of attention will be symbolically mediated and thus the task becomes more automatic (Jonides, 1980; Müller & Rabbitt, 1989; Remington, Johnston, & Yantis, 1992), that is, does not rely on cognitive resources. Results showed no SF effects on cueing effects. Because SF effects only occur when cognitive resources are needed, Wühr and Huestegge (2010) assumed that PO consumes cognitive resources, which leads to narrowing attention, as Baron (1986) postulated with the Overload Hypothesis.

Garcia-Marques, Fernandes, Fonseca, and Prada (2015a)

In this study, the researchers addressed increased monitoring of undesirable interferences in PO by using the composite face effect (Young, Hellawell, & Hay, 1987). When individuals must decide whether two halves of a face (upper and lower) belong to the same face, they are more prone to judge that the two halves belong to the same face when the two halves are perfectly aligned than when they are misaligned, even when the two halves do not belong to the same face. This suggests that individuals process faces holistically. Garcia-Marques et al. (2015a) used the composite face effect paradigm by manipulating PO (coaction) and NPO. Results showed lower composite-face effects in PO than in NPO. Because face perception is holistic in nature, these data suggest that the results were caused by a better outline of the discrepancies between the two halves. The authors clarify that participants in PO were not relying more on holistic processing as a "dominant" well-learned response, but instead were increasing their monitoring of the interference produced by those automatic responses.

Wagstaff, Wheatcroft, Cole, Brunas-Wagstaff, Blackmore, and Pilkington (2008)

These authors directly related data that supports the Overload Hypothesis (Baron, 1986) with the interference of PO in general executive functions. However, contrary to previous studies that show that SF increases control over undesirable influences (e.g., better performance in Stroop tasks) the assumption here was that overload reduces levels of control when necessary. The authors assumed that PO leads to more prefrontal cortex activation, resulting in less capacity to perform tasks that require the use of Executive Functions (e.g., more complex tasks and/or tasks that demand switching). Additionally, by demanding more from the prefrontal cortex, PO "frees up" other areas that use more automatic processing—the type of processing used in routine tasks or in cluster formation. These hypotheses were tested with a phonetic task and a semantic task in which switches (which depend on executive control) and cluster sizes (which do not depend on executive control) were measured. In each of the phonetic task trials, participants were asked to write a word that began with a specific target letter (e.g., "F": food, form, or feel). A cluster was a sequence of words in which the two first letters were the same (e.g., "feel, feed, feature" is a cluster of three words) or a sequence of words with similar sounds (e.g., "some, sum"). Switches were the transition between these clusters (e.g., "food, formation, feel, feed"). In each of the semantic task trials, participants were asked to "write down different foods that they could find in a supermarket." A cluster in this case was defined as a sequence of foods belonging to the same category (e.g., "banana, orange" for fruits category), while a switch was a transition between categories (e.g., "banana, orange, kiwi, beef, bacon, beer, wine" = two switches, one between the categories, fruits and meat, and one between the categories, meat and drinks).

Participants performed these two tasks in either an NPO condition or in one of two PO conditions in which they performed the task in coaction: 1) a monitored condition, in which the experimenter stayed in the room while the participants did the task in coaction; or, 2) non-monitored condition, without the presence of the experimenter. Results showed larger clusters and fewer switches in both PO conditions than in NPO, suggesting that PO disturbed the executive functions (switching), but facilitated more automatic processes (clustering). No difference was found between the two types of PO.

Belletier, Davranche, Tellier Dumas, Vidal, Hasbroucq, and Huguet (2015)

This paper addressed how PO affects executive control (measured by performance in the Simon Task; Simon, 1990). Controlling for participants' working memory capacity (WMC; assessed

with the Reading Span Task; Daneman & Carpenter, 1980), NPO condition was contrasted with two PO conditions: the presence of a confederate and the presence of the experimenter. Participants performed the Simon Task (Simon, 1990), in which conflict trials are established between the position of the stimuli and the position of the keys to be pressed when answering (e.g., if a red circle appears on the left side, participants must respond by pushing the red key on the right). Non-conflict trials position the stimulus and the appropriate key on the same side. Results typically show that participants respond more slowly in conflict trials than in non-conflict trials. SF effects were only found when NPO was compared with the PO condition in which the experimenter was in the room (which suggests no effects of mere presence, but actual evaluation); this condition led to more Simon Interference, suggesting an executive control impairment. Moreover, different relationships were detected between Simon Effects (i.e., interference) and individuals' WMC depending on the condition: a negative relationship in NPO, no relationship in the confederate's presence, and a positive relationship in the presence of the experimenter. However, these positive relationships found within participants' WMC and the Simon Effect were observed exclusively among the slower responses. This was because top-down suppression of incorrect automatic responses are built up only during longer RTs (Ridderinkhof, 2002), which indicates the use of WMC to cope with the Simon Interference. Thus, executive control functions were not impaired in alone or mere presence conditions, but only in the presence of the experimenter and only for those with high WMC (for which a better executive control should be observed; Engle, 2002). A possible explanation for this is that participants with higher WMC attend better to the distractions coming from the presence of the experimenter in the room. Despite the main results of this study being related to the presence of the experimenter and not just mere presence, they give an important picture of how overload can be built into social contexts, and suggest—perhaps counterintuitively—that participants with greater WM are the ones more prone to being affected by overload. Additionally, there is an important difference between the mere presence condition and the NPO condition that was neglected: while working memory is negatively associated with the Simon Effect in NPO, in the mere presence condition there is no such relationship. This can be interpreted as evidence that individuals in NPO have more free cognitive resources and can used them to cope with the conflict (i.e., Simon Interference), while in mere presence, as evidenced by the lack of relationship between Simon Interference and WM, individuals are apparently not using the full potential of their WM capacities, possibly because mere presence itself causes overload.

Belletier and Camos (2018)

This paper addressed the question of whether the foregoing attentional capture by the experimenter's presence affects participants' working memory. According to the Overload Hypothesis (Baron, 1986), social presence is expected to deplete attention and damage memory performance. While the authors did not find effects of PO in memory performance, in a second experiment they showed that the presence of the experimenter in the room interfered with the role that articulatory rehearsal had in improving memory performance. The suppression of articulatory rehearsal was manipulated by asking participants during their task to repeat the sounds "ba-bi-bou"; results showed that suppression was more detrimental in the presence of the experimenter than in the NPO condition. Despite not being directly tested, these results suggest that participants use different memory processes in PO versus NPO; in PO, the use of articulatory rehearsal may play a key role, but much less so in NPO. Additionally, these results once more support the hypothesis that the social presence of the experimenter (which is arguably different from mere presence or coaction) automatically captures attention, which in turn is more likely to diminish cognitive control.

Steinborn and Huestegge (2019)

In this study, Steinborn and Huestegge (2019) asked individuals to do arithmetic problems (i.e., verify if the sum of particular numbers was correct) in PO or NPO. Their results revealed that individuals in PO were faster at solving the arithmetic problems, even the complex ones. Despite this, they did not find any effect in the number of errors. They argue that their results favor the position that PO leads to better processing in general, which is counter to the position that PO consumes attentional resources. However, their arithmetic problems were too simple (e.g., 2+6 = 9), including the problems they defined as complex (e.g., 2+6+8 = 18). That is, despite the increased complexity of summing three numbers rather than two, it is debatable whether this difference is enough to claim arithmetic complexity, given that individuals do these types of arithmetic sums in everyday situations (e.g., shopping, using coins to make purchases, and calculating change). It would be more relevant to operationalize complex problems using multiplication problems, for instance. To summarize, there is a strong possibility that Steinborn and Huestegge's (2019) results are only applicable to easy tasks, which does not permit the elimination of the attentional overload hypothesis in PO when the tasks are legitimately more complex.

The foregoing review of papers focusing on executive control functions and narrowing attention does not clarify how SF depends on them. Results reviewed above show SF effects in decreasing Stroop interference and face composite effects, which suggests an increase in executive control in PO, without excluding the possibility of narrowing attention. However, subsequent studies have argued the opposite. When analysis is focused only on pure PO conditions, the conclusions are that PO increases switching costs (Wagstaff et al., 2008) and has no effect in the Simon task (Belletier et al., 2015), despite affecting the way that WMC is used to cope with Simon Interference.

One common problem with studies that use the "executive control function" term, is that they do not specify what "function" is, as if executive control functions as a control with many mechanisms. This vision of Executive Control as a general control makes it difficult to understand whether the narrowing attention mechanism is also an executive control mechanism, since the Stroop Task is defined in the literature both as a measure of narrowing attention and an executive control function (e.g., Agnew & Agnew, 1963; Ozonoff & Jensen, 1999); notably, Augustinova and Ferrand (2012) treat the two concepts as distinct. Similar to the problem with the term "arousal" (Blascovich et al., 1999), the executive control construct seems to also suffer from unspecified definitions, as the term was born to define different types of functions, which have in common the fact that they regulate one's thoughts and direct behavior toward a general goal (Baddeley & Hitch, 1974). However, tasks differ in what type of control they require. Executive control can be divided into three function categories: updating (i.e., monitoring and tracking working memory information), shifting (i.e., the cognitive ability to switch between tasks or mental sets), and inhibition (i.e., the ability to control interference) (see Miyake & Friedman, 2012; Miyake et al., 2000). The tasks used in the reviewed studies likely operationalized different functions; for instance, Simon tasks focus a participant's capacity to monitor and control for an undesirable interference (inhibition), while fluency tasks are dependent upon different functions (Shao, Janse, Visser, & Meyer, 2014). Thus, performance in verbal fluency tasks may reflect other functions and mechanisms such as updating (Shao et al., 2014), working memory (Henry & Crawford, 2004; Rosen & Engle, 1997; Rende, Ramsberger, & Miyake, 2002), and inhibition (Hirshorn & Thompson-Schill, 2006). This problem with the definition of Executive Control Function, when added to PO conditions with experimenter presence (i.e., evaluation scenario) or mere presence manipulations, creates obstacles to understanding which mechanisms are really underlying SF effects, because it accounts for PO itself and not evaluation effects. Yet, research that

directly studies the narrowing attention mechanism seems to agree that both mere presence and other conditions (e.g., social comparison and evaluation) rely on narrowing attention.

Another approach to SF effects, beyond executive control or narrowing attention mechanisms, involves individuals' context-sensitivity to the tasks. This effect is typically described as an integration of contextual features and the task performance in PO, as illustrated in the studies reviewed below.

Fonseca and Garcia-Marques (2013)

As previously stated, the first studies in SF, developed by Allport (1920, 1924), showed that participants make increasingly more objective associations with a target stimulus when in PO than when alone. Fonseca and Garcia-Marques (2013) proposed that this indicates that participants in PO are attending more to every contextual detail and thus are sensitive to the overall context. The authors offer support for this assumption with two studies. They first replicated the Allport (1920) free association task, but controlled for associative commonality (i.e., how the concepts are socially shared in an associative network to different degrees; Collins & Loftus, 1975; Anderson, 1983). Participants performed the associative tasks either alone in a room (NPO) or in one of two PO conditions: 1) coaction, where two participants did the task at the same time, or 2) mere presence, where the participants did the task in the presence of a confederate who was reading a magazine. The results replicated the Allport studies (i.e., participants made more associations in PO than in NPO). Additionally, the results showed that the associations made in the two PO conditions extended further across the associative network than those made in the alone condition. In sum, participants in PO used more context-related information because their associative networks are wider, which suggests that in PO participants are more sensitive to all possible associations related to the relevant target.

The results of the second study, which used the framed-line test (Kitayama, Duffy, Kawamura, & Larsen, 2003), further supported the hypothesis that participants in PO are more context sensitive. In the framed-line test, participants are presented with a square frame with a vertical line printed inside, extending downward from the center of the upper edge of the square. Next, participants are presented with an empty frame of a different size and asked to draw a line: a) identical to the line in the first frame in absolute length (absolute task), or b) proportional to the height of the surrounding frame (i.e., accounting for the frame size differences; relative task). For the absolute task, a participant only has to attend to the line inside the first frame, and so the frame around the first line is irrelevant for the task. For the relative task, a participant must attend to and incorporate the size of the first frame to draw a proportional line (i.e., they have to incorporate the context). Results revealed that

participants in PO performed better; that is, compared with the NPO condition, there was less difference between the original line and the drawn line, but only in the relative tasks. There were no differences in absolute tasks.

Other studies have also shown evidence of greater sensitivity to context in PO, even without this framework.

Thomas, Skitka, Christen, and Jurgena (2002)

These authors showed that participants' evaluations of an experimenter were more sensitive to cues offered by the social context when in PO than when in NPO. Participants in PO (presence of two confederates) or NPO, were asked to evaluate the experimenter (on the basis of a past interaction), who employed a variety of either positive nonverbal cues (e.g., smiling, positive eye contact, pleasant tone of voice, and handshake) or negative nonverbal cues (e.g., scowling, poor eye contact, and impatient foot tapping). Compared with participants in NPO, those in PO rated positive experimenters more positively and negative experimenters more negatively. This suggests that participants integrate more contextual features into their evaluation process when in PO than when in NPO.

Puntoni and Tavassoli (2007)

This paper presents two studies with SF manipulations: a Lexical Decision Task and an Image Recall Task. Results indicated that words and concepts related to social desirability (e.g., beauty, charm, cute, funny, liked, and lovely) were more activated and more easily recalled in PO than in NPO. Puntoni and Tavassoli (2007) interpreted their data as evidence that social desirability is automatically activated in PO (i.e., people are concerned about the impressions that others are forming about them). However, another explanation for this data is that PO simply sensitizes the mind to social demands (i.e., makes participants more sensitive to the features of the social context).

Garcia-Marques, Fernandes, Prada, Fonseca, and Hagá (2015b)

To show that participants in PO are more context-sensitive, these authors used the Ebbinghaus Illusion task (Doherty, Tsuji, & Phillips, 2008). The Ebbinghaus Illusion effect is the impression that a target circle is larger (smaller) than it really is when it is surrounded by other smaller (larger) circles. Thus, it is an effect that depends on context (i.e., the size of the surrounding circles). In the Ebbinghaus Illusion Task, participants are presented in each trial with two target circles and their respective surrounding circles, which may be larger or smaller. The participant's task is to identify

the largest target circle. Results revealed that participants in PO performed worse than those in NPO, which suggests that it was more difficult to ignore undesirable contextual influences (i.e., the context) in that condition. Arguably, then, PO demands more context-sensitive processing.

Summary

For many years, the literature has confounded SF effects with several other effects that can be generated in PO. For this reason, many biased statements and assumptions relating to SF have been made. Therefore, the focus of this analysis is only on effects that can be attributed to SF, and not on others such as evaluation and social comparison.

The new evidence presented suggests that in PO, participants more fully integrate all contextual features into their task performance, which is revealed through content (e.g., associations used), performance facilitation (e.g., drawing a line by accounting for the frame context), and/or performance impairment (e.g., by strengthening contextual illusions). Following Fonseca and Garcia-Marques' (2013) view, these data suggest the possibility that in PO our minds allow for more spreading of activation.

However, as all of Huguet's (Huguet et al., 1999; Sharma et al., 2010) studies show, PO also modulates executive control functions, increasing control over undesirable influences, possibly by narrowing the individual's attention. Thus, while it can be assumed that more information is available in PO, it must also be assumed that individuals are more selective in PO. In fact, when combining the evidence for greater spread of activation and of narrowing attention in PO, it is clear that some incompatibilities can arise. However, these incompatibilities may not be real. An increase in spreading activation in PO may itself impose a need for an increase in narrowing of attention, as suggested by Baron (1986). The increase in available information leads participants to attend only to what is central to the task and less to peripheral cues (Easterbrook, 1959; Cohen, 1978). However, this may be more likely when the available information is undesirable or unnecessary for performing the task. Thus, by "activating" more peripheral cues, the increase in spreading activation can be expected to result in better performances when the cues directly assist the task goal (e.g., Fonseca & Garcia-Marques, 2013), and in worse performances when the cues do not help the task (e.g., contextual illusions; Garcia-Marques et al., 2015b).

Therefore, this thesis accounts for the possibility that two mechanisms are increased in PO: one related to the likelihood that associated information is activated (spreading activation), and

another related to the control processes that filter which information should be attended/ignored (narrowing attention). As such, SF likely results from the interplay between the two mechanisms.

Activation and control are basic common cognitive processes that are claimed to characterize many social cognitive phenomena (e.g., Fazio, 1986; Brewer, 1988; Fiske & Neuberg, 1990), including stereotyping. Stereotypes are knowledge structures that can be activated by context cues, influencing our behavior when no control is exerted to prevent their influence. Stereotyping, then, is likely influenced by PO, which makes it possible to address a general hypothesis about the ways in which SF influences the interplay of control and activation mechanisms.

In the next chapter this hypothesis is addressed by reviewing available literature about SF effects on Stereotyping.

Chapter II
Stereotyping and effects of Social Facilitation

To clarify why SF effects can be expected with regard to stereotyping, this chapter reviews the literature on this connection. First, stereotypes are defined, and how their impact on our behavior can be inferred (how it can be measured). Next, stereotype activation is defined and conditions are discussed that favor its activation and expression, and how this activation can be influenced by SF. Then, possible dynamic features between stereotype activation and stereotype control processes (i.e., whether activation is controllable or stereotypes are controlled only in their expression) are discussed, with the aim of clarifying how these processes can be modulated by SF. Finally, the available empirical studies showing that SF has a key role in stereotyping are introduced.

Stereotypes as structures of knowledge, and their activation

The term stereotype in Social Psychology was first introduced by Lippmann (1922) as "pictures in our heads". This vague definition gained some specificity with studies showing that when individuals thought about specific groups such as jews (Katz & Braly, 1933) or "hooligans" (Macrae, Stangor, & Milne, 1994), they consistently labeled their association with specific traits: mercenary for jews, and rude for hooligans. These cognitive knowledge associations between people's intergroup representations and specific traits is what is commonly referred to as stereotypes in the scientific literature (see Fiske & Taylor, 2013).

Allport's (1954) definition of a stereotype was, "an exaggerated belief associated with a category." However, as Kunda (1999) noted, more recent stereotype definitions did not focus on the degree to which the associations were exaggerated or accurate. Rather, they focused only on the type of associations linked to a social group; that is, knowledge, beliefs, expectations, feelings, and thoughts (see Hamilton & Sherman, 1994; Esses, Haddock, & Zanna, 1993; Fazio, Jackson, Dunton, & Williams, 1995). Empirical evidence shows that these associations of specific traits with specific groups, (e.g., aggressiveness with black people) repeatedly bias our decisions and judgments (see Fiske & Taylor, 2013). For instance, in the who-said-what experimental paradigm (where statements are

associated with different members of a group and participants are asked later to match statements with individuals), participants make more memory errors related to confusing people within categories (e.g., age, race, gender, or sexual orientation) than between categories (Taylor, Fiske, Etcoff, & Rudermam, 1978; Maddox & Chase, 2004; Rakić, Steffens, & Mummendey, 2011). Graham and Lowery (2004) also showed that the same crime description resulted in worse punishments when the perpetrator was described as black than when they were described as white. Macrae, Milne, and Bodenshausen (1994) illustrated, moreover, that it is easier in an impression formation task to infer that the person is intelligent, shy, and short if they are labeled as an Asian American. After exposing individuals to an ambiguous behavior (i.e., a person being shoved), Duncan (1976) showed that if the actor was a white person, their behavior was judged as playful; however, if they were a black person, they were evaluated as aggressive.

Following Bodenhausen, Todd, and Becker (2006), stereotyping studies can be split into two approaches. One is a bottom-up approach, which looks at why and how stereotypes are shaped. This approach postulates that people are continuously forming impressions or making inferences about others as a means of interpreting their social world. Thus, they learn about categories of individuals by forming knowledge structures (i.e., categorization) that summarize and generalize the traits they identified in such social groups (i.e., stereotypes). The other approach focuses on top-down mechanisms, that is, on how those knowledge structures (stereotypes) support or bias our behaviors/judgments. The aim is to understand how the social category, activated in one's mind on the basis of specific person-features (e.g., gender or skin color), can support trait and behavior expectations about an individual (e.g., judging black-skinned individuals' behavior as aggressive) and modulate one's judgments.

The work of Devine (1989) was a milestone in stereotype research—it showed the universal nature of stereotypes in their impact on cognition and behavior. In her first experiment, Devine asked her participants to write characteristics relating to black people, based not on their own personal beliefs but on their cultural knowledge. Her results showed that characteristics such as poor, aggressive, criminal, and low intelligence were consistently given, independent of the participant's prejudice level, as measured by the Modern Racism Scale (McConahay, 1986). This suggests that stereotype structures can be strongly shared across participants, at least when they are asked directly. However, are stereotype structures activated when participants are not questioned directly about them?

This led Devine (1989) to set up a second experiment, in which she subliminally exposed her participants to words generally associated with black people (e.g., negroes, lazy, black, blues, rhythm, and unemployed). Although these words are not directly related to the concept of hostility, the participants who were primed with them judged ambiguous behaviors as more hostile compared with those who were not primed. These results show that stereotype structures are not only shared but also activated in the presence of stereotype cues (i.e., words/concepts that are already part of the stereotype structure). Bargh, Chen, and Burrows (1996) and Chen and Bargh (1997) gave further support to Devine's conclusion by showing that subliminal priming with actual black portraits, and not words, also led participants to judge others as more hostile and ruder. Clearly, the first two experiments (Devine, 1989) illustrate that stereotype structures are shared and activated in the presence of concepts/cues associated with the stereotype, leading to biased judgments, even when those cues are outside of our awareness. So, are we all biased by stereotypes?

The third experiment (Devine, 1989) seemed to indicate that we are not all biased by stereotypes. As in experiment 1, Devine asked participants to write down cultural associations about the black stereotype. However, after providing cultural associations, they were asked to do a thought-listing task, in which they wrote down their personal thoughts about black people (personal thoughts). The results revealed that high- and low-prejudice participants, as measured by the Modern Racism Scale (McConahay, 1986), did not differ in terms of word valence on the first task (i.e., number of positive and negative associations). This indicates equal stereotype activation and valence independence. However, on the personal thought-listing task, high-prejudice participants revealed a higher frequency of negative thoughts about black people than low-prejudice participants. Thus, despite equal stereotype knowledge across all participants, those less prejudiced somehow avoided expressing negative concepts associated with the stereotype. This suggests that there is an option to not use the stereotype (i.e., cultural associations and negative-black association); in other words, a person can choose to not attend to it.

Devine's (1989) findings reinforced the idea that although stereotypical information is available to us, we can decide to attend to other information provided by the target (e.g., individual or social group). This is an assumption made by Fiske and Neuberg's (1990) Continuum Model of Impression Formation. The authors proposed that the processes used to form an impression/opinion about others follow a continuum between category-related information and person-specific information. Category-related information is assumed to rely on few cognitive resources, to be unintentional, and

to be based on salient category/stereotype cues (e.g., skin color, ethnicity, age, or gender). Most importantly, category-related information is assumed to have priority over person-specific information. This led to the use in this theory of person-specific information only if the perceivers see the other as relevant (i.e., motivational goals). In a scenario in which an individual is not perceived as relevant, impression formation only uses category/stereotype-related information. In a scenario in which an individual is perceived as relevant, the perceiver is motivated to allocate cognitive resources to person-specific information. If the person-specific information is consistent with the category/stereotype, the perceiver's final impression will be based on that category/stereotype. However, if the person-specific information is inconsistent with the initial categorization, the perceiver will try to recategorize the target to form a suitable category, which can lead to the personalization of the target.

Devine's (1989) studies support the assumptions of the Continuum Model of Impression Formation (Fiske & Neuber, 1990; see Brewer, 1988, for an alternative view); they show that stereotype content is strongly shared, and when activated, it colors our judgments. However, only when motivation and capacity resources are available do people proactively attend more to all available information (individuated and categorical information). The general idea is that we need cognitive resources to avoid the impact of a stereotype that was automatically activated (via allocation of our attention to individual information). Although this general assumption is highly consensual (see, for instance, Macrae & Bodenhausen, 2000), there is also evidence in the literature showing that stereotype activation may be more likely in some groups than in others (e.g., depending on individuals' prejudice levels; see Lepore & Brown, 1997), and that stereotype activation itself may also rely on cognitive resources (see Gilbert & Hixon, 1991). Thus, the impact of reduced cognitive resources may rely on whether the stereotype, once actived, will support further processing (Gilbert & Hixon, 1991; Govorun & Payne, 2006; Sherman, Macrae, & Bodenhausen, 2000).

Stereotype activation

Although SF effects can be expected in stereotyping, it is not clear whether they occur through stereotype activation. There are, however, several ways in which this might occur. One route relates to the overload condition created by the presence of others (PO); this condition may either reduce individuation or facilitate it by narrowing participants' attention to it (Cohen, 1978; Baron, 1986). This phenomenon can be disturbed by a lack of cognitive resources, and thus connects the use of stereotype information to the cognitive demands of individuating (Macrae et al., 1994) and also to category

activation itself. Gilbert and Hixon (1991) provided evidence of this by using a word completion task: they presented participants with fragmented words such as S_Y or POLI_E, and asked them to choose a letter to complete the word (e.g., SKY or POLICE). The authors demonstrated that when participants had an Asian assistant, they used more Asian stereotype words to complete the fragments (e.g., SHY and POLITE) than when they had a non-Asian assistant. This suggests that the presence of the Asian assistant worked as a cue to activate the stereotype, which gave participants access to specific stereotype associations. However, when participants were asked to rehearse an eight-digit number during the task (which engaged cognitive resources), the effect disappeared. The authors concluded that cognitive overload disturbs stereotype activation, which further suggests that without cognitive resources some cues cannot activate the stereotype.

Thus, the expected SF effects may depend on the assumption that stereotype activation is automatic and, therefore, resource independent (see Bargh, 1989; Brewer, 1988; Fiske & Neuberg, 1990; Freeman & Ambady, 2011; Kunda & Thagard, 1996) or on the assumption that category activation requires some cognitive resources (Gilbert & Hixon, 1991). It should be noted, however, that even if cognitive overload clearly prevented stereotype activation in Gilbert and Hixon's (1991) approach, it was not clear how the activation was prevented. One possibility is that stereotype activation relies on cognitive resources, which were used for the number rehearsal rather than for the activation of stereotype associations. Another possibility is that the cognitive overload reduced attention to external cues (i.e., assistant's race), and so there was no stereotype activation because attention was focused on the main tasks and not on the stereotype cue. These two scenarios are vastly different and have different implications for the understanding of how stereotypes can be modulated by social contexts (i.e., coinciding or not with Baron's narrowing attention hypothesis; Baron, 1986). Macrae, Hewstone, and Griffiths (1993) suggest it is more likely that cognitive overload causes narrowing attention (i.e., decreasing attention to external cues) than that it steals resources that could be used for stereotype activation. In a different paradigm, Macrae et al. (1993) asked their participants to watch a video of a female doctor or a female hairdresser interacting with another person. During the interaction, the female doctor/hairdresser displayed some behaviors consistent with their profession while others did not. For instance, the hairdresser expressed enjoyment of disco, while the doctor liked attending the opera. While watching the video, half of the participants were in cognitive overload (rehearsing an eight-digit number), while the other half was not. After the video ended, all participants were asked to recall the interaction they had observed on the videotape. Results showed that information consistent with the doctor/hairdresser stereotype was unchanged by the load

manipulations, while more inconsistent information was recalled in the no overload condition than in the overload condition. This seems to indicate that consistent stereotype information is like a heuristic that does not need to active cognitive resources. However, inconsistent information depends on cognitive resources, which are reduced when participants are in overload. Note that in this study it was important to focus on the video actors (and their characteristics), as it was essential to attend to their conversation. However, in Gilbert and Hixon's (1991) study, the main task was to complete fragmented words; the video actors were merely assistants who gave the instructions (i.e., not very relevant for the task). These experimental differences explain why in Macrae et al. (1993) stereotypes were still being activated in overload. Because stereotype information is relevant for the task, a narrowing attention via cognitive load does not draw individuals' attention to that stereotype information. Conversely, in Gilbert and Hixon (1991), stereotype information was an external cue, so narrowing attention prevented stereotype activation by drawing individuals' attention to stereotype information. These two studies (Gilbert & Hixon, 1991; Macrae et al., 1993) are relevant to an understanding of how SF effects can affect stereotype activation. The previous chapter showed that an SF context could be understood as an overload context that can disrupt highly demanding processes (Wagstaff et al., 2008) or can lead to a narrowing of an individuals' attention (Baron, 1986). Therefore, SF should not prevent stereotype activation in tasks in which stereotype cues are intrinsic to the stimuli to be attended to.

Spencer, Fein, Wolfe, Fong, and Duinn (1998), in a close replication of Gilbert and Hixon (1991), also presented a relation between environment cues (external cues) and stereotype activation. Like Gilbert and Hixon (1991), Spencer et al. (1998) used the Word Completion Paradigm with overload manipulation but added a feedback manipulation. Their feedback manipulation was based on Sinclair and Kunda (1999), who observed more stereotype activation when participants received negative feedback from a black manager than when they received the same feedback from a white manager. These results tend to be interpreted as evidence that self-threat motivates individuals to search in the environment for cues to protect themselves (e.g., the use of stereotypes to discredit negative opinions). Thus, Spencer et al. (1998) added a bogus intelligence test before the Word Fragment Task; the feedback to their test results was manipulated (positive vs negative). Results revealed that those who received positive feedback related it to their intelligence, which replicated the overload effect of Gilbert and Hixon (1991) by showing no stereotype activation in the overload condition. However, participants who received negative feedback activated the Asian stereotype, even in the overload condition. This pattern of results was also obtained in experiments 2 and 3, but by

using subliminal primes of the black stereotype before the word fragment tasks. The authors interpreted these results as suggesting that negative feedback (i.e., self-esteem threat) motivated participants to search for reasons to make downward comparisons; thus, they searched for stereotypical information that would help discredit the negative opinions, even in overload conditions where cognitive resources are scarce.

The literature continues to view stereotype activation as a phenomenon independent of cognitive resources. Perhaps this is because Gilbert and Hixon's (1991) results are now believed to be caused by the narrowing of attention, which does not allow the cue to automatically activate the stereotype, rather than as evidence that stereotype activation is not an automatic process (see Sherman, Macrae, & Bondenhausen, 2000). However, Macrae, Bondenhausen, and Milne (1995) have shown that two different stereotypes cannot be activated at the same time, which suggests that stereotypes compete for activation. This conclusion was reached within a paradigm in which their participants were subliminally primed with the words "Chinese" or "woman" before watching a movie that depicted a Chinese woman reading a book. Their idea was to show that by priming one stereotype, the other would not be used with the Chinese woman categorization. Results revealed that participants primed with "Chinese" more quickly identified words related to the Chinese stereotype (e.g., trustworthy and calm), while participants primed with "woman" more quickly identified words related to the woman stereotype (e.g., emotional and romantic). This facilitation induced by the primed category was also followed by a restraining of the non-primed category; that is, participants were slower than the control group (no priming) to identify words associated with the non-primed category. These results suggest that, while stereotypes are automatic in nature, they seem to compete for activation, as if a kind of limited resource pool exists that is specific for the activation of stereotype content.

In sum, the dependence of stereotype activation on cognitive resources is still debatable, given that stereotype activation occurs easily but also relies on a specific and limited resource pool that does not permit simultaneous activation of two stereotypes. Nevertheless, when in overload, individuals seem to narrow their attention, which allows them to prevent stereotype activation only if their attention is diverted from peripheral stereotype cues. If the cue is directly related to the task, activation can be expected. These assumptions are essential for understanding how SF can modulate stereotyping when the task makes the stereotype a central cue. In these conditions, PO should increase stereotype activation either by increasing attention to relevant cues or by imposing cognitive demands.

Importantly for this thesis, when stereotype cues are central to the task, and a stereotype is already activated in one's mind, its expression seems to be stronger in overload, which suggests that there is also a need for resources to prevent the expression of the stereotype after its activation. This kind of mechanisms are reviewed below.

Stereotype expression: Inevitable or controllable?

Despite the automatic feature of stereotype activation, their influence is not inevitable if stereotype cues are attended to. As Devine (1989) illustrated, low-prejudice individuals can avoid responding in a stereotypical way even when they have knowledge of the stereotype. This result seems to indicate that, although stereotypes can be activated, individuals can respond in ways that are not influenced by it. This suggests that either the activation can be suppressed or that, even after its activation, the stereotype expression can be avoided.

An example of the control of stereotype activation is given by Macrae, Bodenhausen, Milne, and Jetten (1994). They established a paradigm in which their participants wrote an essay about a day in the life of a skinhead. However, half of the participants were specifically asked to not use stereotype information in their essay (suppression condition). The participants who had been given suppression instructions expressed less stereotypical information than those participants who had not been given suppression instructions. However, when participants were asked to write another paragraph about another skinhead, this time without suppression instructions, the results were inverted. Participants who had previously suppressed stereotype information used more stereotype information in their essays (study 1). This result suggests that the use of suppression as a means of not being influenced by stereotype information activation—despite being initially effective—leads to subsequent hyperaccessible activation of the stereotype. The authors also replicated this hyperaccessibility using behavioral measures; in study 2, after writing the paragraph, participants were told, falsely, that they would meet a "skinhead" in another room. The room that the participants were taken to had many chairs, one of which held some of the "skinhead's" belongings, suggesting that the "skinhead" had temporarily left the room. While "waiting" for the "skinhead's" return, the participants were instructed to choose a seat, whereas those in suppression group were the ones to seat farther away from the "skinhead's" chair in comparison with those in the no suppression group. The results of study 3, in which participants performed a Lexical Decision Task after writing the paragraph, also revealed a hyperaccesibility pattern—those in the suppression condition were faster to identify words associated with a skinhead stereotype. It is important to note that this type of suppression does not avoid

stereotype activation; rather, participants who are asked to suppress stereotype subsequently show more stereotype activation and expression.

As Devine's (1989) experiments suggest, stereotype activation does not necessarily lead to stereotype expression. Although stereotype activation occurred for all participants, which seemed to be culturally shared, its expression depended on participants' prejudice levels. Research by Sherman, Stroessner, Loftus, and Deguzman (1997) clarify that even in the suppression of stereotype expression there is activation. The authors asked their participants to watch a video of an Asian female presenting reasons why she should be accepted to a particular school. Ten of the reasons were stereotype related such as, "does everything her boss tells her to do even when she doesn't think it is her responsibility," while the other ten reasons were stereotype unrelated such as, "makes dinner at home." The authors also manipulated suppression by asking half of the participants to form an impression about the person without using stereotypes. After watching the videotape, all the participants completed a recognition task; the results showed higher recognition of stereotypical behaviors by participants in the suppression condition than for the other participants. This seems to indicate that when individuals are asked to not express stereotypes, they start to monitor for possible stereotype information so they can decide to not show/express it. This monitoring, then, must activate stereotype content to inhibit its influence (i.e., to suppress a specific thing, we must search for/focus on it), which leads us to activate the stereotype.

The control exerted over stereotype expression seems to also occur when there is no instruction to do so. For instance, by measuring participants' event-related potentials (ERPs) during a sequential prime task (Payne, 2001; see next section), Amodio, Harmon-Jones, Devine, Curtin, Hartley, and Covert (2004) detected higher activation of the anterior cingulate cortex (brain region associated with conflict detection; see Berns, Cohen, & Mintun, 1997; Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001) in incongruent stereotypical trials (e.g., Tool primed by a Black portrait) than in congruent trials (e.g., Gun primed by Black portrait). More interesting is that this higher activation of the anterior cingulate cortex also occurred for correct responses, which indicates that even when responding correctly, participants were activating stereotypes whose influence was somehow inhibited and thus was not expressed. Moreover, this inhibition seems to have a cost—correct responses in incongruent trials were slower than those in the congruent trials. This indicates that the inhibition of stereotype influences imposes extra effort compared with when inhibition is not necessary.

These studies suggest that even if stereotype activation occurs, its expression in our responses/judgments is not inevitable. We can avoid stereotype expression not only by avoiding its activation (e.g., Gilbert & Hixon, 1991) but also by avoiding its influence in our responses, which implies monitoring the effects of its activation (e.g., Amodio, et al., 2004). This capacity to avoid stereotype interferences in our judgments is what the stereotype literature refers to as Stereotype Control.

Because SF effects are assumed to impact control mechanisms (see Chapter I), those mechanisms are another avenue through which PO can interfere with stereotype expression. In the next section, it is reviewed how activation and control mechanisms interplay in stereotyping, and how SF can modulate it.

How stereotype activation and control interplay has been studied.

On the basis of their experiments on attention and perceptual searching, Shiffrin and Schneider (1977) suggested that these processes differ in their automaticity, that is, they can be either automatic or controlled. Automatic processes are broadly defined as effortless, and inevitable through the presence of a triggering stimulus that activates a sequence of memory nodes, hence, influencing perception, judgment, and/or action in a determined way. Controlled processes use a temporary sequence of nodes to complete a specific task. Because these nodes are not activated spontaneously by the presence of a trigger, some kind of "effort" is necessary to keep the nodes activated. Thus, controlled processes are constrained and dependent on cognitive resources.

The simultaneous engagement of controlled and automatic processes in stereotyping gained great relevance after Devine's (1989) conclusions. At the same time, researchers began to realize that because of the role of control processes, it was sometimes difficult to understand the effects of stereotype activation. The use of explicit measures (e.g., surveys and interviews) to test stereotype activation was put at risk; these methods are highly prone to social desirability, which make possible his respondents to use control mechanisms in order to not be prejudiced. This did not offer suitable conditions for studying automatic stereotyping features. Thus, many researchers not only relied more on implicit measures but also created new ones (see Gawronski, 2009).

The origins of the term "implicit" suggest that such tasks/measures are free of social desirability distortions, and so capture only what is more purely automatic. However, as Ito et al. (2015) pointed out, "implicit" does not mean that the measures cannot be influenced by effortful/control processes. In fact, many researchers (e.g., Conrey, Sherman, Gawronski, Hugenberg,

& Groom, 2005; Jacoby, 1991; Payne 2001, 2005) have illustrated this mutual interplay of processes when developing ways of dissociating automatic and controlled processes in the same task. It is thus essential to understand the nature of the experimental task, as some afford a better understanding of one process or another. Below is a review of the most commonly used implicit tasks in the stereotyping field, as well as how they measure automatic and control features.

Pronunciation and Lexical Decision Tasks

Of all the implicit measures presented in this chapter, Pronunciation and Lexical Decision Tasks (LDTs) seem to be most strongly related to purely automatic processes. The rationale for these tasks is to make individuals decide whether a target set of letters is or is not a word (e.g., "chair" would be a word, but "rhcia" would be a nonword). However, the addition of stereotype primes (e.g., photo of a woman) prior to viewing target words that are related (e.g., sexy or weak) or unrelated to the stereotype (e.g., chair and table) leads to faster lexical decisions in the former versus the latter scenario (see Moskowitz, Gollwitzer, Wasel, & Schaal, 1999).

This effect serves as evidence that the presence of a stereotype cue (prime) facilitates the activation of stereotype-related words, which allows the measurement of how activated some concepts are in an individual's mind when they are primed by a specific stereotype. Incidentally, Moskowitz et al. (1999) also used this task to show that individuals who endorse egalitarian norms (i.e., the principle of non-discrimination) do not show this facilitation effect, which suggests that these individuals can somehow control stereotype activation even when the SOA (Stimulus-Onset Asynchrony) is short (200 ms). Because the literature indicates that consciously controlled processes cannot occur when SOAs are below 600 ms (see Bargh, Chaiken, Govender, & Pratto, 1992; Fazio, Sanbonmatsu, Powell, & Kardes, 1986), Moskowitz et al. (1999) claimed that an unconscious mechanism permits those who endorse egalitarian norms to avoid activation of the stereotype.

Later, Moskowitz and collaborators (Moskowitz, Salomon, & Taylor, 2000; Moskowitz & Li, 2011) extended their results by using black portraits as primes and by using egalitarian words as targets in LDTs. Again, those who endorsed egalitarian norms did not display facilitation effects for stereotype words associated with black people; however, they did display facilitation effects for egalitarian words. This suggests that, when presented with a stereotype cue (e.g., a black person's portrait), those who endorsed egalitarian norms activated egalitarian concepts instead of stereotype associates. This restraint of the activation of the black stereotype is very similar to Macrae, Bondenhausen, and Milne's (1995) findings, which showed that two different stereotype concepts can

compete for activation. The research of Moskowitz and collaborators demonstrated that egalitarian concepts and stereotype concepts also compete when activation of one restrains the activation of the other. Because these effects happen in short SOAs (Moskowitz et al., 1999) it can be postulated that these tasks are more closely related to automatic activation processes (in which egalitarian concepts and stereotypes share the same pool of activation resources; see Macrae et al., 1995), as it is impossible to study more traditionally high-order control processes in such tasks (e.g., Shiffrin & Schneider, 1977).

Implicit Association Test

Another task that was initially used to measure stereotype association activation is the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998). The main idea of the IAT is to create specific blocks of trials in which participants' responses can be influenced by associations in their mind. The first IAT to measure stereotypes activation consisted of five blocks of trials. In the first block, participants identify names typical of white people (e.g., Meredith) or of black people (e.g., Latonya). Thus, in each trial, a name appears at the center of the screen, while two category labels appear at the top of the screen: Black at top right and White at top left. The positions of the categories must be attended to in order to give the correct response (i.e., to choose Black, participants must press the right key; to choose White, they must press the left key). In the second block, instead of identifying names as typical of black/white people, participants evaluate words (e.g., lucky or disaster) as Pleasant or Unpleasant. These first two blocks were created only to familiarize participants with the response keys and the labels. In the third block, four categories appear at the top instead of two: two on the right (Black and Pleasant) and two on the left (White and Unpleasant). Again, participants must judge names as Black or White and non-name words as Pleasant or Unpleasant. Because Black and Pleasant are on the same side, this block is called incongruent or counterstereotypical. Typically, Unpleasant—a negatively valence classification—has been associated with the Black Stereotype. The fourth block repeats the first block, but the labels are reversed (i.e., White at the top right of the screen and Black at the top left). Finally, the fifth block presents the four labels again, but this time with White-Pleasant at the top right screen position and Black-Unpleasant at the top left, which makes this block congruent with the stereotypical association, black-unpleasant.

The use of congruent and incongruent blocks in the same task creates two ideals but different situations: one in which stereotype activation facilitates performance (congruent block) and another in which stereotype activation leads to worse performance (incongruent block). In this way, people

with stronger black-unpleasant associations (i.e., stereotype) would arguably respond faster in the congruent block than in the incongruent block, thus revealing their implicit negative attitude towards black people.

Greenwald et al. (1998) tested the IAT with categories such as flowers-insects, instruments-weapons, Asian names versus American names, and, as mentioned, typical names of black and white people. Their results showed that, in general, white participants had faster RTs for white-pleasant combinations than for black-pleasant combinations, which indicates a bias toward seeing white as more pleasant than black. Additionally, their data demonstrated that for flower-insect categories, the IAT strongly correlates with explicit measures about insects and flowers (i.e., feeling thermometer and semantic differential); however, the other IAT versions (Asian-American and Black-White), in which participants may have motives for hiding their prejudice, the correlation between the IAT and explicit measures is weak. This suggests that the IAT can be immune to self-presentation motivations, and thus also to control processes that can hide stereotype activation. Although IAT data can be analyzed by simply comparing the RTs of incongruent and congruent blocks, it is common to use an IAT *d* index as a measure of association between categories, which also penalizes error responses (see Greenwald, Nosek, and Banaji, 2003).

Greenwald et al. (1998) assume the IAT is highly sensitive to stereotype activation but immune to control processes. However, their repeated use of the task soon showed that was not the case. For instance, the application of a Multinomial Model to the IAT data (a mathematical approach, based on Error Rates (ERs), to estimate different processes; see Batchelder & Riefer, 1999; Riefer & Batchelder, 1988) made it clear that the influence of controlled processes is lower in incongruent blocks (where stereotype activation and control processes oppose each other) than in congruent blocks (where stereotype activation and control processes contribute to the same response), and can be affected by control manipulations such time restriction (Conrey et al., 2005).

Weapon Identification Task

The Weapon Identification Task (WIT; Payne, 2001) is a sequential priming task in which participants have to decide whether they saw a handgun or a tool. However, in each trial, before the presentation of the target (handgun or tool), participants are primed by a black or white portrait. An important feature of the WIT is brief (SOA 100 ms) presentation of the target, which forces

participants to respond without the opportunity to confirm what they saw. Judd, Blair, and Chapleau (2004) adapted the original WIT to also offer evidence of a positive racial bias effect by using the association of black people with basketball. This suggests that sequential priming tasks are also effective with positive/non-negative black stereotype associations (e.g., strong, sport, and rhythmic), and not only with negative ones (e.g., dangerous or threatening).

The use of stereotypical trials (Black-Gun) and counterstereotypical trials (Black-Tool) makes this task ideal for capturing the stereotype association, black as criminal/dangerous (Devine & Elliot, 1995). It is also possible in this task to model activation and control components by comparing congruent and incongruent trials. One such model is the Process Dissociation Procedure (PDP; Jacoby, 1991; Jacoby, Toth, & Yonelinas, 1993).

The PDP relies on the fact that the WIT involves trials in which control and automatic processes influence individuals' responses in the same direction (i.e., congruent trials) and in opposite directions. Thus, a Black prime followed by a Gun target presents participants with a congruent situation because both controlled perceptual processing (Control Component) and stereotype association between Black and Guns (Automatic Component) lead to the same responses. This allows for the statistical computation of: 1) a control (C) component, because of the difference between choosing Gun in congruent versus incongruent trials (C = Congruent-Incongruent), and 2) an automatic (A) component, because of the probability of choosing Gun in incongruent trials when C fails (A = Incongruent/(1-C)).

The results of Payne's (2001) experiment 1 showed that participants primed by black (white) portraits were quicker to correctly identify guns (tools) than when they were primed by white (black) portraits, as component A was higher for Black Primes than for White primes. This indicates that in Black trials, compared with White trials, participants have a stronger tendency or bias to choose Gun, which indicates a Black-Gun association activation (i.e., a stereotype association). However, this effect was not found for the C component.

To test how restricted time would affect participants' responses to a WIT, Payne (2001) made a version of the task in experiment 2 in which participants had only 550 ms to respond. By adding time restriction, he made the effect disappear in RTs and appear in ERs. PDP data showed again that component A was higher for Black primes than for White primes; in reality, the level of component A was no different from that in experiment 1. What this indicates is that time restriction did not affect automatic bias, which reinforces the notion that stereotype activation is independent of cognitive

resources. As in experiment 1, there was no difference for component C between Black and White primes. However, compared with experiment 1, levels of component C were lower, which suggests that the time restriction caused participants to use controlled perceptual processes less accurately. Because component C was equal in White and Black primes, even though component A was higher for black than for white primes, it is suggested that the control processes captured by the PDP were exerted equally across all primes (i.e., a more general type of control). However, the higher level of component A observed in Black trials versus White trials, suggests that when control fails, participants are more likely to choose Gun in black trials than in White trials (i.e., they are influenced by stereotype activation, but only when control fails).

The idea that biased misidentifications in WITs are caused by general control failures was tested by Payne, Shimizu, and Jacoby (2005). They contrasted two possible explanations for WIT-biased misidentifications: a) Illusory Perception hypothesis, in which participants believe they are seeing a different object because of the cueing effect of the stereotype prime, b) Executive Control Failure hypothesis, in which participants fail to control their responses even when their perception of the target is not altered by the prime. To test which mechanism was more likely to happen, they applied the WIT with a time restriction, but in this case, after each response, participants had to rate their confidence about their responses. If the bias effect was caused by illusory perception, participants would be expected to have equal confidence in their errors and their correct responses. However, in the case of executive failure, participants would have a higher confidence rating for their correct responses than for their incorrect responses. Results showed that participants accurately evaluated their correct target identifications, and also their misidentifications; in other words, participants had higher confidence ratings after giving a correct response than after giving an incorrect response. Note that this effect is unrelated to stereotype consistency (i.e., primes). A second experiment was set up; this time, instead of asking for confidence ratings after their first responses, participants were asked to respond a second time but without time limits. This alteration gave participants the opportunity to correct their responses. Results showed the stereotype pattern when accounting only for first-time responses. However, the pattern disappeared after accounting for second responses, which indicates that participants could monitor their errors very well. Putting these results together, it appears that the effects on the WIT only occur when participants have executive failures. Payne et al. (2005) concluded that stereotypes influence participants' judgments/actions when executive functions (i.e., control) fail, and not when target perception is distorted. Govorun and Payne (2006) strengthened this claim: to exhaust participants' cognitive resources before a WIT, half of their participants did 300 Stroop trials

(depletion condition) while a control group did just 30 (no depletion condition). This manipulation caused depletion participants to commit more stereotype errors than the non-depleted participants, which again suggests that the executive functions are crucial to stereotype expression.

Payne (2005) also conducted a related study in which he compared the WIT with other tasks and explicit measures. His data revealed that the C component of the WIT correlated with a concern for control (explicit measure), while the A component related to performance on evaluation tasks, such that individuals with higher levels of A made stronger associations between negative words and black primes. This seems to indicate that the WIT captures two different processes, one related to controlling responses in a way that achieves the right judgment, and another related to automatic evaluations about the race of the prime.

Curiously, the stereotype index derived from the WIT is not related to the stereotype index derived from performance in the IAT. This suggests that the two tasks do not capture stereotype activation and stereotype expression in the same way. The interplay between the components of the different processes can be different in these two stereotypical tasks.

The description of tasks presented in this section clarifies that they capture not only stereotype activation but also control processes, which are assumed to prevent participants' responses from being biased by the activated stereotype. Thus, it is clear that to detect any modulation by POs of task performance, researchers must focus on the interplay between the two components: one on activation and another on control over possible bias effects. Above, it has been stressed how PO can be expected to modulate stereotype activation. Here, it must be stressed that if social presence is assumed to enhance control over interference introduced in a Stroop task (e.g., Sharma et al., 2010), it could also reduce the level of interference that the activated stereotype exerts over participants' behaviors. However, if social presence is assumed to reduce cognitive resources and to reduce other possible relevant executive control functions (Wagstaff et al., 2008), control should arguably be weaker in PO.

Fortunately, there are studies that have already tried to address SF effects on stereotyping, and which have allowed alternative hypotheses to be conceptually tested. In the next section, some of these studies are reviewed and the means by which SF effects are thought to influence some of the experimental tasks are presented.

Research on Social Facilitation effects on stereotyping

Only two papers have directly addressed SF effects on stereotyping: Lambert, Payne, Jacoby, Shaffer, Chasteen, and Khan (2003), and Castelli and Tomelleri (2008). Their results and conclusions are, however, not congruent with each other. Following is a detailed critique of their evidence and results, which highlights the aspects of their work that informed the present approach.

Lambert, Payne, Jacoby, Shaffer, Chasteen, and Khan (2003)

Lambert et al. (2003) were the first to propose exploring the effects of SF on stereotyping. They aimed to test the hypothesis that individuals in a public context reveal more prejudiced responses compared with those in private contexts. In their first experiment, Lambert et al. (2003) accessed participants' attitudes towards black people at two time points. First, all the participants' attitudes were accessed (anonymously) merely by a set of questionnaires. At the second time point, two months later, participants completed an Impression Formation Task in the laboratory. However, half of the participants were told that their responses would be confidential (private context), while the others (public context) were told, "after you have formed your impressions of this person, there will be a general discussion session with the other participants in the room today. During this discussion, each of you will have the opportunity to show others the information that you were given, as well as talk about the kinds of judgments you made about this person." The Impression Formation Task consisted of the presentation of a biographical sketch about Donald, a black individual. After the presentation, participants responded to a set of questions intended to access their attitudes (e.g., "how much would you want meet this person?"). In the end, all participants also replied to a set of anxiety questionnaires.

Results showed that the congruency between participants' black attitudes at point 1 (accessed by questionnaires) and point 2 (accessed by the impression formation task) was directly related to their anxiety levels (i.e., more anxiety led to more attitude congruency). However, this only happened for participants in the public context. While participants in the private context did not differ in terms of anxiety levels compared with those in the public accountability condition, no relation was detected between their attitude congruency and their reported anxiety. Lambert et al. (2003) interpreted these results as evidence that arousal in public contexts leads to a greater use of dominant responses; in other words, the use of internal attitudes in impression formation tasks.

To test whether this effect occurred due a greater activation of stereotypical content (via arousal), or due a decrease in controlled processes (via narrowing attention; here interpreted as depletion of cognitive resources), the authors conducted a second experiment. Employing the same

manipulations—private vs public contexts—they used the WIT (Payne, 2001) for their experimental task. Results were typical for WIT tasks (the version with time restriction): higher ERs in Black-Tool (White-Gun) trials than in White-Tool (Black-Gun) trials. As expected, this pattern was stronger for participants in the public context (supposedly the PO condition) than for those in the private condition. This suggests that in public-anticipation settings, participants evidence a stronger bias. Using PDP estimates (Jacoby, 1991), Lambert et al. (2003) found no difference in Automatic components promoted by context manipulation. However, they detected lower Control component levels (for both black and white primes) for participants in the public context than for those in the private condition. This pattern again suggests the use of a general type of control in the WIT (Payne, 2001; Payne et al., 2005) to not allow stereotype activation and, thus, bias. In contrast to Experiment 1, anxiety levels were higher for those in the public context than for those in the private context.

Unfortunately, there are several problems with the interpretation of SF effects on stereotyping in the results of this study. First, as stated in Chapter I, the authors did not use an actual SF manipulation; specifically, they used anticipated public context manipulations (accountability): participants were told that their responses would or would not be discussed at the end of the experimental session. The authors' justification for this type of manipulation was: "…because social facilitation effects are not restricted to cases in which the audience is physically present, as such findings can arise even when participants anticipate or imagine that others might be appraising their work." (p. 279). Their argument demonstrates that their understanding of SF effects (behavior in PO vs NPO) was entangled with accountability effects (behavior with vs without anonymity). Another critical caveat in their approach is the possibility that there was no real PO condition in their experiments, that is, their participants did the tasks in an individual booth, regardless of their accountability condition (with or without anonymity). This restricts their results to a comparison between the effects of NPO with accountability and NPO without accountability, as there was no PO condition that could test the actual SF effect (i.e., mere presence).

Nevertheless, this study provides a relevant experimental setting in which to study SF effects on stereotyping. More importantly, it suggests the use of PDP components to disentangle the mechanism beyond the effect.

Castelli and Tomelleri (2008)

Castelli and Tomelleri (2008) do not connect SF with the typical theories of SF effects presented in Chapter I. Their rationale for this study of SF effects on stereotyping was to show that PO induces people to pay more attention to egalitarian norms, which postulate equality between humans. So, they expected responses from participants in PO versus those in NPO to be less stereotype-biased. SF conditions were manipulated, leading participants to perform the task "either alone or in the same room with two peers" (Castelli & Tomelleri, 2008, p. 2).

In their experiment 1, they used the IAT to measure implicit attitudes towards black people and two items to access the presence of individuals' egalitarian norms (i.e., "It is understandable if a company decides to hire a White rather than a Black worker even though they have similar curricula"; "It is understandable that a girl on a train prefers to sit beside a White rather than a Black male"). Their results showed higher *d600* indexes for participants in NPO than for participants in PO, which indicates stronger negative attitudes towards black people in NPO than in PO. No differences between the SF conditions were observed with regard to egalitarian values. However, a negative relation between the norm and the *d600* index was observed in PO, while no relation was detected in NPO, which suggests that individuals' responses were influenced by egalitarian norms only in the PO condition.

In a second experiment, the authors attempted to illustrate that egalitarian norms are more activated in PO than in NPO. For this, they used a version of the LDT similar to that used by Moskowitz et al. (2000): before each word/nonword participants were primed with Black/White portraits. Castelli and Tomelleri (2008) hoped that trials primed by Black portraits would facilitate the choice of words rooted in egalitarian concepts. They used egalitarian words (e.g., tolerance and equality) as targets, and positive words irrelevant to egalitarian norms (e.g., kindness and responsibility) as control words. The results showed that, in general, egalitarian words more quickly identified than non-egalitarian words. However, in PO, this effect was moderated by the type of prime: with black primes, egalitarian words were chosen faster than non-egalitarian words, while for white primes, RTs for egalitarian and non-egalitarian words were equal. Participants in NPO, moreover, showed only the main effect of faster RTs for egalitarian words than for non-egalitarian words, independently of the prime. A summary of the LDT results (experiment 2) suggests that in PO the presence of a black race cue (prime) leads to activation of egalitarian concepts, which explains why participants in PO were less prejudiced when doing an IAT (experiment 1).

In contrast to Lambert et al. (2003), Castelli and Tomelleri (2008) used actual SF manipulations with two different tasks. While it is suggested that using an IAT produces less stereotype bias in PO, the authors showed, by using a LDT, that this effect can be caused by increased activation of egalitarian norms in PO. The use of actual SF manipulations and the convergence of the results in two different tasks is a strong argument for believing that SF in stereotyping leads to less bias. However, Castelli and Tomelleri (2008) did not thoroughly explore how activation of egalitarian norms affected participants' performances in the IAT. Does the activation of egalitarian norms cause participants to expend more effort in PO to avoid showing prejudice? Or, is the activation of egalitarian norms constrained during stereotype activation? The authors view the latter hypothesis as more plausible; like Moskowitz and colleagues (Moskowitz et al., 1999; Moskowitz et al., 2000; Moskowitz & Li, 2011), they believe that the activation of egalitarian norms competes with the activation of stereotype associations, and thus when one is activated, the other is constrained. However, it would be ideal if Castelli and Tomelleri (2008) had provided more support for this hypothesis through their IAT data (e.g., RTs, ERs, and PDP components), but they only reported the *d600* index data.

How can these data improve our understanding of SF in stereotyping?

The findings of these two studies (Lambert et al., 2003, Castelli & Tomelleri, 2008) claim that PO has different effects on stereotyping: PO either increases or decreases stereotyping. Interestingly, given that PO can modulate the interplay between activation and control, both of these two possibilities can be expected.

The two studies show that the effect is dependent upon two different mechanisms: Lambert et al. (2003) posit that the increase in stereotyping is related to the fact that individuals in PO reduce their control over the influence of the activated stereotype. This suggests that PO induces a general overload, which hinders only the control over stereotype expression.

Castelli and Tomelleri (2008) propose that the activation of egalitarian values can overcome stereotype activation (either by suppression or by activation of concurrent information), and thus stereotypes are less activated in PO.

How can these apparent contradictory effects be explained?

First, it should be noted that, as strange as it may appear, the two studies share a common finding: the use of a "dominant response." In Lambert et al. (2003), experiment 1, individuals' performances in public contexts (which the authors assume is a PO condition) relied more on their available response, the stereotype. Similarly, Castelli and Tomelleri (2008; experiment 1) showed that it is in PO that individuals' egalitarian norms are predictive of lower-prejudice responses. This suggests that individuals in PO draw more on their own inner states (e.g., beliefs, attitudes, and values).

Second, the two sets of studies used two different tasks, which the literature asserts are not related (Payne, 2005; Ito et al., 2015). As such, the modulation by PO may depend on the processes that dissociated the two tasks.

Finally, the two studies may not have captured the same social presence conditions. Only Castelli and Tomelleri (2008) had an isolated condition compared with a co-action condition. Because Lambert et al. (2003) did not manipulate actual SF conditions, it is difficult to know whether their results would replicate under such conditions.

To address the relevance of all of these differences, the present research/thesis aimed to replicate the two sets of studies using similar experimental settings. Replicating the procedures used by Lambert et al. (2003)—the WIT in actual SF conditions (PO vs NPO)—this thesis attempts to clarify whether PO effects match those "accountability effects." One possibility is that by using NPO conditions and a co-action condition, the results will follow the same direction as those reported by Castelli and Tomelleri (2008). However, if the results replicate Lambert et al. (2003), even with actual SF conditions, then it will be important to understand whether the different results relied on the task itself: either the IAT or the WIT (as suggested by Payne, 2005). The next chapter addresses these issues.

Chapter III
The effects of Social facilitation on Stereotyping: When replication reveals the unexpected

This chapter describes the empirical approach taken in this thesis to the study of Social Facilitation (SF) effects on stereotyping. First, the SF effect on stereotyping is defined; individuals exhibit different levels of stereotyping depending on whether they are alone or in PO. Then, the cognitive mechanisms beyond stereotyping are disentangled.

The empirical strategy was to replicate the two studies that have identified different stereotype effects (Lambert et al., 2003; Castelli & Tomelleri, 2008) within the same social context; that is, a context in which social presence was manipulated using a co-action setting. One important distinction between these two studies was overcome by approaching SF effects as differences detected between participants in NPO (being alone in absence of the experimenter) versus PO (being in co-action with others). Furthermore, this approach avoided attributing the effects to factors such as evaluation, social comparison, and accountability. It is important to stress the need for excluding the presence of the experimenter in the NPO condition since that presence can be perceived as evaluative by the participant (see Chapter I).

The empirical work conducted for the purposes of this thesis went further than mere replications of the effect, and aimed to also disentangle the cognitive mechanisms that are related to stereotype expression: activation and control. As such, Lambert et al. (2003) were followed and the two components were addressed using the PDP approach (Jacoby, 1991). As in the original studies, the tasks were the WIT and the IAT. Both have conditions where control and automatic processes lead to the same response (congruent) and conditions where those processes lead to different responses (incongruent). Thus, both in WIT trials where the prime is a Black portrait and the Target is a Gun, and in the congruent IAT block where the Black and Negative categories share the same response key, the stereotype activation and control processes lead to the same response that facilitates participants' performances. However, when the Prime is a Black portrait and the Target is a Tool (in

the WIT), or the categories Black and Positive share the same response key (in the IAT), the Stereotype activation leads to the wrong response, while the control processes help them choose the right responses. As Jacoby (1991) ingeniously noted, this permits the creation of equations that not only disengage the influences of automatic and control processes on participants' performance (accuracy) but also help estimate their influences through two indices (C and A). For instance, in congruent trials where Control (C) and Automatic (A) processes lead to the correct response, the participant's likelihood of choosing the correct response can be denoted as equal to $C + A(1\text{-}C)$. However, in incongruent trials, the two processes lead to different responses. Thus, the probability of an error is statistically represented in incongruent trials as $A(1\text{-}C)$, that is, Automatic influences when Control processes have failed (or are not present).

In this way, the PDP can be applied to accuracy data by assuming that the influence of Control (C) on performance is equal to Correct responses in congruent trials minus Incorrect responses in incongruent trials, while A can be calculated by dividing error rates for incongruent trials by 1-C (i.e., when Control is not present). Having established these two process indexes, statistical tests were conducted to determine how SF effects rely on activation and control processes. If data replicated Lambert et al. (2003), it could be expected that SF effects occur because of less control being exerted in PO (as observed in their public condition).

Further, to capture stereotype effects across RT distributions, we followed Sharma et al. (2010), and thus expected to find evidence of increased control over time in PO, when compared with NPO. This type of approach has its beginning with Ratcliff (1979), who argued that effects that are analyzed only with RT averages (or measures of central tendency) tend to ignore their distribution (i.e., how the effect behaves throughout RT distributions). These authors proposed analyzing differences between experimental conditions, not only on the basis of individuals' central tendencies but also across their own RT distribution. To accomplish this, it was necessary to analyze participants' Cumulative Distribution Frequency (CDF) plots—these represent small response intervals (Bins) that are ordered from faster to slower intervals. Further, each Bin has an effect index (e.g., Stroop interference, calculated by subtracting the RTs of Incongruent trials from those of the Congruent trials for each bin), which allowed the behavior of the effect to be mapped across each participant's accumulated RTs.

Mapping how the effect behaves throughout the participants' RTs was elucidative for the aims of this thesis, because, as in Sharma et al. (2010), the Stroop effect increased over the participants'

RTs; however, this effect decreased during participants' slower responses in PO (i.e., the last bins), suggesting that they were controlling the Stroop interference.

Bins that represent Stereotype indexes, or Delta plots (because the bin to capture the effect is the result of the difference between two other bins, Incongruent minus Congruent), can help us not only understand whether the Stereotype effect increases across RTs, which was verified in other interference tasks beyond Stroop (e.g., Ridderinkhof, Wildenberg, Wijnen, & Burle, 2004; Davranche, Hall, & McMorris, 2009; Burle, Spieser, Servant, & Servant, & Hasbroucq, 2014), but also understand whether Social Conditions (PO and NPO) cause differences in how the Stereotype effect behaves across participants' RTs. For instance, as in Sharma et al. (2010), it is possible that in PO the Stereotype effect is better controlled in slower responses.

In sum, evidence of stereotyping offered by the IAT and the WIT allowed us to test the hypotheses that SF occurs because PO increases Stereotype bias (Lambert et al., 2003) or because it decreases Stereotype bias (Castelli & Tomelleri, 2008). PDP and CDF analyses, moreover, helped us understand why PO, or NPO, promotes such biases.

Why might PO increase stereotyping?

1) Because in PO people are more sensitive to contextual features, which increase mental activation (e.g., Fonseca & Garcia-Marques, 2013), it may cause them to: a) attend more to peripheral stereotype cues, and/or b) experience quicker stereotype activation.

2) Because, following Zajonc's (1965) Drive Theory, stereotypes are assumed to be dominant responses (Lambert et al., 2003), and dominant responses increase in PO compared with NPO contexts.

3) Because PO is an overload context, in which some relevant executive control functions are challenged (Wagstaff et al., 2008). This may lead to stronger stereotype bias in PO by decreasing the efficiency with which an individual can cope with stereotype activation (as found by Lambert et al., 2003).

Why might PO decrease stereotyping?

1) Because PO is an overload condition that causes individuals to narrow their attention to the relevant features of the task (e.g., Cohen, 1978; Baron, 1986). This, in turn, can lead them to attend

less to possible contextual/peripheral stereotype cues, which reduces the probability of stereotype information activation.

2) Because in PO individuals deal more efficiently with interferences (Huguet et al., 1999; Sharma et al., 2010), so even if interference is stronger, time will allow them to overcome it more efficiently.

3) Because, as also suggested in the SF literature, executive control functions are better executed in PO (e.g., Muller, et al., 2004; Augustinova & Ferrand, 2012). Thus, individuals in PO are more capable of dealing with/controlling stereotype information.

4) Because in PO it is easier to activate relevant contextual features (e.g., Fonseca & Garcia-Marques, 2013) such as egalitarian concepts (Moskowitz et al., 2000; Castelli & Tomelleri, 2008).

5) Because of the more efficient interplay between activation of conflict information and its correction. There is evidence in both stereotype literature (Amodio et al, 2004) and SF literature of this inhibitory process, which seems to be stronger in PO (see Garcia-Marques et al., 2015b).

All these possibilities suggest not only that PO can modulate stereotyping processes differently, but also that the role of PO is highly complex; this may explain why the two available studies produced contrary results. The empirical approach taken in this thesis offers further clarification of the foregoing possibilities.

Experiment 1- Weapon Identification Task (WIT) with actual SF conditions

The main aim of this study was to replicate the results of Lambert et al. (2003) with the WIT, but with actual SF conditions (PO and NPO). It was expected that, if these results fully replicated Lambert et al. (2003), there would be less control (indices) in PO than in NPO.

Experiment 1 - Methods

Participants and Design

A total of 136 non-black ISPA (Lisbon, Portugal) undergraduates (21 men, mean age: 21.99 years) volunteered to participate in the experiment and were randomly assigned to a 2 Social Conditions (NPO vs PO) x 2 Prime (Black vs White portrait) x 2 Target (Tool vs Gun) design; the first condition was a between-subjects factor, while the remaining conditions were within-subjects

factors. The sample size was determined following Westfall, Kenny, and Judd (2014); a minimum sample size was determined for a power of .80 based on effect sizes calculated from Lambert et al.'s (2003) experiment 2 data (stereotyping effect, *d*=.85; and its moderation by Accountability, *d*=.46) and the number of trials (384 trials).

Procedure

Participants gave their informed consent to participate in a lab experiment that used face and object identification. No reference was made to race or social setting manipulations. Experimental sessions were scheduled as a group or individually to manipulate the social conditions. Individual participants (NPO) were accompanied by the researcher to a laboratory room and left alone in front of a computer screen to complete the task. Grouped participants (PO) arrived in sets of 6 to 10 and were assigned to individual computers to perform their tasks at the same time (i.e., co-action). In the PO condition, the experimenter remained in the room. However, to reduce possible evaluation effects, the experimenter worked on their computer (inactive attention) during the sessions and sat at a table positioned in a way that made it impossible for them to monitor participants' computer screens.

The experimental instructions, which were displayed on the computer screens with *E-prim*e 2 Software support (Psychology Software Tools, Pittsburgh, PA), explained that the task was to identify an object presented on the screen as either a Tool (with the letter L of the keyboard, right side) or a Gun (with the letter S of the keyboard, left side). They were told that each object would be briefly preceded by a portrait; however, only the objects were to be identified as a Tool or a Gun, and should be identified correctly and as fast as possible.

To familiarize themselves with the targets and the response time window, participants performed two practice blocks. First, they were shown 32 objects sequentially and asked to identify them as either a Gun or a Tool; no time-restriction was imposed. Second, they were shown 40 objects sequentially in intervals of 550 ms and asked to identify them as either a Gun or a Tool.

Once a participant or group completed their practice blocks, they began the WIT. This task involved the same objects and identification process, but each object was preceded by a prime of either a Black or White portrait. Both Prime and Target pictures were 5.3 cm x 4 cm. Primes included portraits of four Black and four White male portraits. Lambert et al. (2003) also used female portraits; however, because Conrey et al. (2005) suggest that the Black-Gun association is only true for male portraits, the traditional version with only male primes was used (Payne, 2001; 2006; Payne et al.,

2005). Targets were four different handguns and four different tools. Like Lambert et al. (2003), each trial began with a visual pattern mask (500 ms), followed by the prime (200 ms), which was replaced by the target (100 ms). After the target, another visual pattern appeared (450 ms), depicted in Figure 1. Thus, each trial provided participants with a response window of 550 ms (target plus the last visual pattern). When no response occurred within the response window, a red exclamation point appeared on the screen (500 ms) before the next trial. The idea was to warn the participant that they had not responded quickly enough. The WIT had a total of 384 trials, with 128 trials per block.

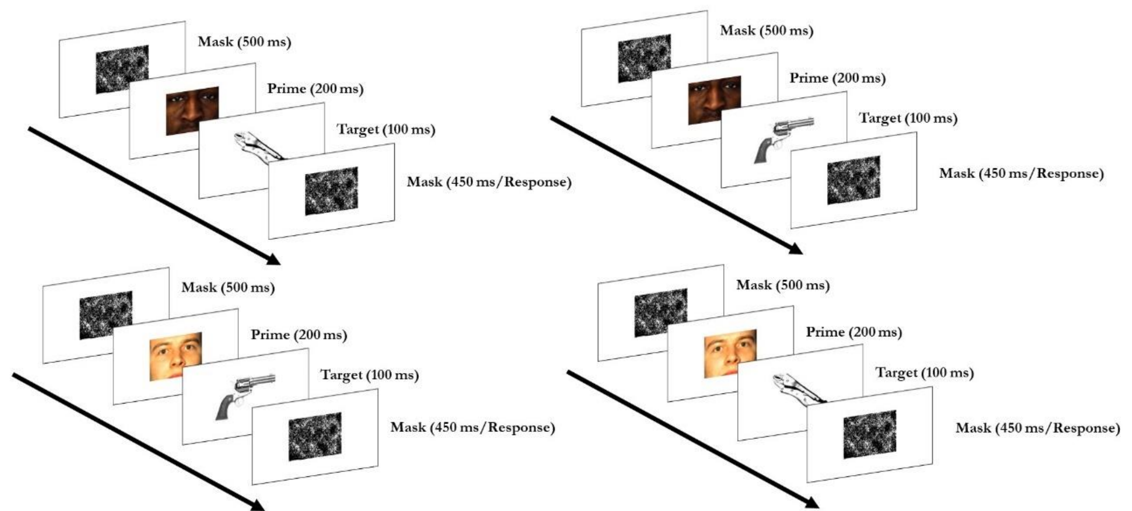

Figure 1. *Weapon Identification Trials. Left trials represent Incongruent trials, where Black(White) primes precede Tools(Guns) target. Right trials represent Congruent Trials, where Black(White) primes precede Guns(Tools) target.*

Upon completing the three WIT blocks, participants completed self-report measures; these were designed to test whether the SF conditions were promoting other effects (e.g., evaluation) than mere presence, and also to access additional evidence for possible SF mechanisms (e.g., motivation and attention; see Table 1).

Table 1

Self-Report Measures

Measure	Items	Rating Scale Extremes
Engagement	I think in this study I engage...	Nothing/Much
Motivation	I think in this study my motivation was...	None/Total
Attention	I think in this study my attention was...	None/Total
Difficulty	I think that the tasks of this study were...	Very Difficult/Very Easy
Capacity	I think that the tasks of this study demanded...	Little/Much of my Capacities
Accompaniment	During the task I felt...	Alone/Accompaniment
Observation	During the task I felt...	Unobserved/Very Observed
Evaluation	During the task I felt...	Unevaluated/Very Evaluated
Positive-Negative	How do you feel at this moment?	Positive/Negative
Sad-Happy	How do you feel at this moment?	Sad/Happy
Rested-Tired	How do you feel at this moment?	Rested/Tired
Bored-Alerted	How do you feel at this moment?	Bored/Alerted
Well-Bad	How do you feel at this moment?	Well/Bad
Tense-Relaxed	How do you feel at this moment?	Tense/Relaxed

Dependent Variables

Error Rates (ERs) were calculated for each Prime-Target combination, where participants' errors were divided by the total number of completed trials. Note that non-responses were not included in ER calculations.

PDP components (Jacoby, 1991) were computed in accordance with Payne (2001) for access control and automatic indices in the WIT. C component, the likelihood of correctly discriminating the target, and A component, the likelihood of choosing Gun when control fails, were calculated for each prime (C component for Black primes, C component for White primes, A component for Black

primes, and A component for White primes). For Black prime trials, the C component was calculated by the expression: Correct responses in Black-Gun trials (%) – Incorrect responses in Black-Tool trials (%). Likewise, A component was computed as: Incorrect response in Black-Tool Trials (%) divided by (1-C), when control fails. To compare the C(A) Black component with the C(A) White component, the components for White trials were calculated in the same direction. C was computed by the formula: Correct responses in White-Gun trials (%) - Incorrect responses in White-Tool trials (%). Likewise, A component was computed as: Incorrect response in White-Tool trials (%) divided by (1-C).

Reaction Times of Correct Responses (RTs) were computed as an average of correct RTs for each Prime-Target combination (i.e., Black-Tool, Black-Gun, White-Tool, and White-Gun).

CDF plots were created to analyze stereotype bias through RT distributions. To achieve this, the following steps were taken. First, each participant's' experimental RT distributions (i.e., Tool-Black, Gun-Black, Tool-White, and Gun-Black RTs) with their associated responses were sorted and split into five bins (percentiles). In the time-restricted WIT version, the stereotype effect is detected in ERs, not in RTs (Payne, 2001), thus RTs of errors were maintained for each RT distribution. Second, to facilitate CDF plot interpretation, one Stereotype index was calculated based on the ERs, (i.e., *(Black-Tool ERs minus Black-Gun ERs) minus (White-Tool ERs minus Black-Gun ERs)*), within each bin. Hence, each participant's Stereotype index comprised five bins, where higher values indicate a stronger stereotype effect.

Experiment 1- Results

Preliminary analysis of the data showed that participants had difficulty responding within the imposed time limit. Eleven of the participants provided too many non-responses (more than 80% of the trials), so their data were excluded. Data from 125 participants—63 in NPO and 62 in PO—were included in the analysis.

To understand how the SF manipulations worked, we first compared the control measures in each social condition to determine whether participants felt accompanied in PO but not evaluated or observed. Evidence of stereotyping and the stereotyping process was then evaluated, with a focus on the impact the SF manipulations exerted on stereotype indexes. For this, a general linear model approach was followed, which integrated all the experimental conditions. However, given that it was the PO condition in which a replication of the stereotype effects found in the literature were expected to be seen, and since only SF studies tend to have NPO conditions, for the sake of clarity, a simple effect analysis is provided wherever relevant.

Self-reports

The two Social Conditions were contrasted relative to how much participants felt: *accompanied* (M_{PO}= 4.29; *SD*=1.57 vs M_{NPO}=3.71; *SD*=1.86; *t*(122)= -1.90; *p*=.06, *d*=.34), *evaluated* (M_{PO}= 3.77; *SD*=2.06 vs M_{NPO}=4.10; *SD*=2.06; *t*(122)=.87; *p*=.39, *d*=.16), and *observed* (M_{PO}= 2.15; SD=1.53 vs M_{NPO}=1.82; *SD*=1.20; *t*(122)=-1.31; *p*=.19, *d*=.24). Taken together, these measures indicated a tendency for participants to feel more accompanied in PO. More importantly, however, there was no difference between the two conditions regarding how evaluated or observed participants felt.

Also relevant for the aims of this thesis, these results showed that participants felt the task was more *demanding*, in terms of their capacity, in the PO condition (M_{PO}= 5.69; *SD*=1.14) than in the NPO condition (M_{NPO}=5.15; *SD*= 1.35; *t*(122)=-2.5; *p*=.02, *d*=.43).

The *engagement* (t<1) and *motivation* (t<1) items did not reveal any effects, suggesting that motivational aspects were not influenced by the SF manipulations, as well as self-reported *attention* (*t*(122)=-1.11; *p*=.27, *d*=.20 and *difficulty* (t<1).

The self-reports that are more related with mood, *positive-negative* (t<1), *sad-happy* (t<1) *well-bad* (t<1) and *bored-alerted* (t<1), were not influenced by our SF manipulations.

The self-report item, *rested-tired*, had a marginal effect, *t*(122)=1.75; *p*=.08, *d*=.32, suggesting that participants in PO (*M*=5.47; *SD*=1.32) felt more tired than participants in NPO (*M*=5.00; *SD*=1.64). The self-report item, *tense-relaxed*, however, was not significant *t*(122)=1.75; *p*=.12, *d*=.29.

Error Rates

In the time-restricted WIT versions, stereotype effects were captured in individuals' error rates (ERs; Payne, 2001; Lambert et al., 2003), which had an overall mean proportion of 25%. We further contrasted the ERs of the design conditions by performing a 2(Social Condition) x 2(Prime) x 2(Target) mixed ANOVA (see Figure 2).

Results showed evidence of a general stereotype bias in the WIT responses, as documented by the significant interaction of Target x Prime, *F*(1, 123)=22.59, p<.001, η^2=.16. As expected, participants misidentified more Tools (i.e., confounded them with Guns) when primed by Black faces (*M*=.26, *SD*=.15) than when primed by White faces (*M*=.22, *SD*=.13). The reverse pattern occurred for Guns, which were misidentified more often when primed by White (*M*= .25, *SD*=.14) than by Black faces (*M*= .23, *SD*= .12).

The isolated main effect of the Prime was also significant, F (1, 123) =7.90, p<.01, η^2=.06, indicating higher ERs when participants were primed by Black faces (M=.24; SD=.12) than when primed by White faces (M=.23; SD=.13); this may have been a direct consequence of the Stereotype effect. However, the isolated effect of the Target was not significant (F<1).

In contrast to Lambert et al. (2003), no statistical evidence was found to suggest that the Stereotype effect was moderated by Social Conditions, Target x Prime x Social (F<1). Indeed, a non-reliable main effect of Social Condition, F (1, 123) =2.97, p=.09, η^2=.02, was found, which indicates a tendency for higher ERs in PO (M_{PO}= .26, SD=.12) than in NPO (M_{NPO}= .22, SD=.12).

No other effects (Target x Social Condition; Prime x Social Condition interactions) were significant (F<1).

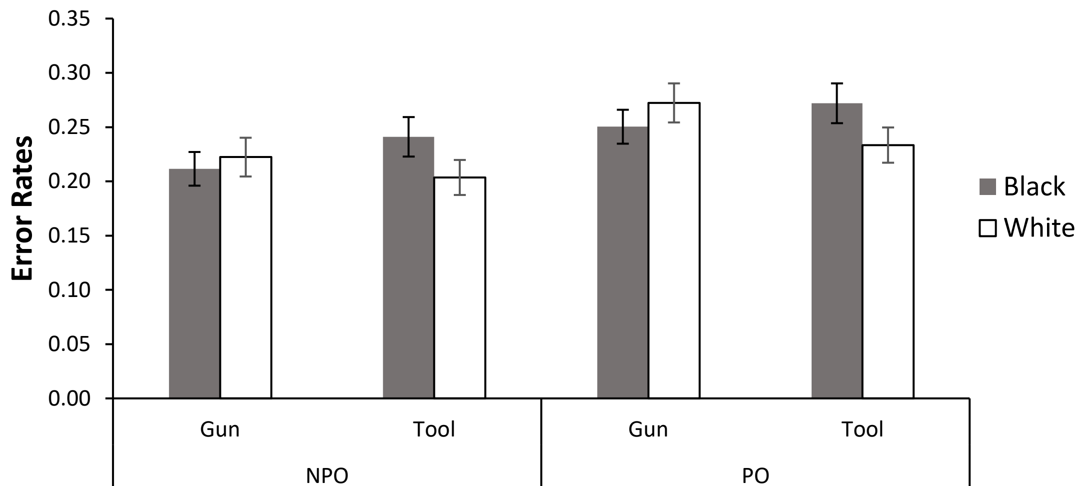

Figure 2. *Error Rates for target (Gun vs Tool) and Prime (Black vs White) WIT trials in both NPO and PO. Errors bars denote one standard error around the mean.*

PDP Components

Following other WIT studies (Payne, 2001; Lambert et al., 2003), estimates of each PDP component for both Black and White trials were computed for each participant (see components estimates and its standard deviations, Table 2). The aim was to address any evidence suggesting that Social Conditions reduced participants' control levels in their responses, as Lambert et al. (2003) found.

The analysis containing the *C component* (likelihood of discriminating tools from guns) as dependent measure in a mixed ANOVA, defined by Prime (Black vs White) and Social Condition

(NPO vs PO) as factors, showed only a non-reliable effect of Social Condition, *F* (1,123) =2.97, *p*=.09, η^2=.02. The direction of this effect matched that of Lambert et al. (2003): less discriminability in PO (*M*=.49; *SD*=.24) than in NPO (*M*=.56; *SD*=.24). Results showed no evidence that the type of control that decreases in PO was associated with any type of prime, as the Social Condition x Prime interaction was not significant, *F* (1,123) =.07, *p*=.79). Thus, if anything, PO merely reduced a general type of control.

Unexpectedly, the main effect of Prime was significant, *F* (1,123) =7.90, *p*<.01, η^2=.06, suggesting that there was higher discriminability in White (*M*=.53; *SD*=.25) than in Black trials (*M*=.51; *SD*=.24). Note that this result was not found in Lambert et al. (2003), nor in Payne (2001). While in their studies, Primes seemed not to affect discriminability, in the present study, Black primes led to less discriminability. Therefore, to verify whether the Prime effect on the *C Component* still presented in the PO condition (which is typical in WIT studies), a simple analysis was performed. This analysis revealed that the Prime effect was not significant in PO, *t*(123) =1.51, *p*=.13, *d*=.13; however, the same simple analysis in NPO revealed the Prime effect, *t((*123) =2.46, *p*=.02, *d*=.21.

The mixed ANOVA (Prime x Social Condition) for the *A component* (i.e., likelihood of choosing Gun when discriminability fails (control)), revealed only the highly significant main effect of Prime, *F* (1,123) =14.77, *p*<.001, η^2=.11. The effect indicated a stronger likelihood of choosing Gun in Black (*M*=.52; *SD*=.13) versus White trials (*M*=.47; *SD*=.12). This effect replicated previous studies (Payne, 2001; Lambert et al., 2003), showing that Black Primes increased the tendency to respond with Gun, which indicates activation of the stereotype association, Black-Gun. As in Lambert et al. (2003), no effect was promoted by Social Condition, nor did it interact with type of Prime (*F*<1).

Table 2

PDP component estimates (Standard Deviation) for Primes and Social Conditions

Components	Social Condition		Overall
	NPO	*PO*	
C-Black	.55(.21)	.48(.27)	.51(24)
C-White	.57(.22)	.49(.28)	.53(.25)
A-Black	.52(.13)	.52(.14)	.52(.13)
A-White	.48(.14)	.46(.11)	.47(.12)

Correct Reaction Times

While time-restricted WIT versions only detect stereotype effects in ERs (Payne, 2001; Lambert et al., 2003), it is nevertheless important to assess whether Social Condition induces different RTs, given the indications in the literature that individuals in PO respond more quickly than those in NPO (e.g., Triplett, 1898). The answer to this may help explain the null effects with regard to ERs. RTs were analyzed within a 2(Social Condition) x 2(Prime) x 2(Target) mixed factorial ANOVA (see Figure 3).

The Social Condition, F (1, 123)=7.30, p<.01, η^2=.06, main effect corroborated our expectations, indicating that participants in PO (M=392.43; SD=46.24) were faster than those in NPO (M=415.00; SD=46.26).

The Target main effect, F (1, 123) =5.81, p=.02, η^2=.05, also suggested that RTs were slower for Tools (M=405.55; SD=47.23) than for Guns (M=401.88; SD=47.75). Importantly, this effect was qualified by the type of Prime and so, unexpectedly, there was a reliable interaction between Target and Prime, F (1, 123) =8.17, p<.01, η^2=.06, which indicates that there was a stereotype effect in RTs. Post hoc analysis showed differences in the time taken to respond in the Tool trials, t((123) =-2.92, p<.01, d=.49, indicating that participants were slower to correctly identify a Tool target when primed by a Black portrait (M=407.64; SD=48.56) than a White portrait (M=403.54; SD=48.31). No difference between Primes was observed for Guns, t((123) =1.01, p=.31, d=.06.

No other effect was significant, including the main effect of Prime (*F* (1, 123) =1.42, $p<.24$, $\eta^2=.01$, and any interaction with Social Condition (all *Fs* <1).

This evidence of stereotyping in RT data is not typically found in the literature. Thus, by running a simple analysis on NPO and PO data, this effect was explored to see whether these data replicated the results found in the literature. This analysis revealed that replication was indeed likely, such that no significant evidence of the stereotype effect was found when considering only the PO condition, $t(123)=-1.61$, $p=.11$. Thus, introducing the NPO condition seemed to account for the effect, a conclusion that was confirmed when only the data from participants in NPO were considered, *t*((123) =-2.44, $p=.02$. This difference in the simple analysis results suggests that participants in NPO dealt with the WIT differently than those in PO, although the results did not provide clear evidence of this.

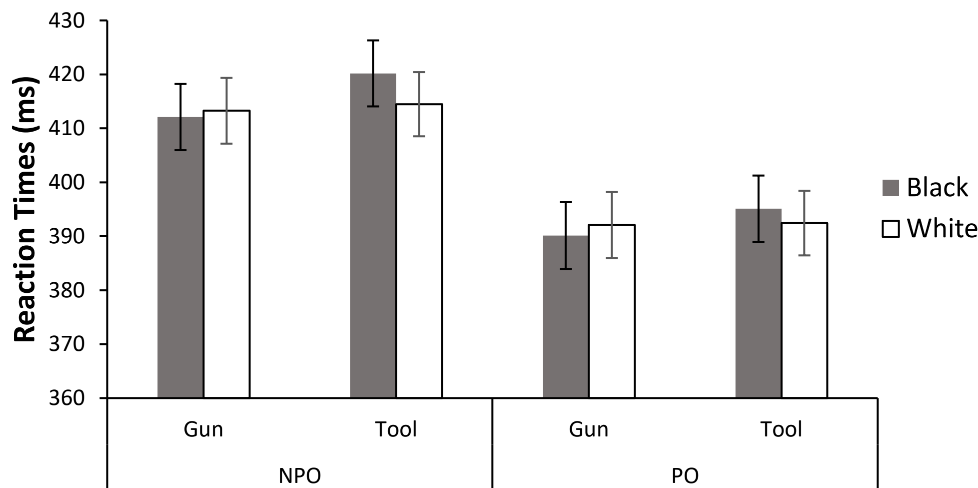

Figure 3. *Correct Reaction Times for Target (Gun vs Tool) and Prime (Black vs White) in WIT trials for both NPO and PO. Error bars denote one standard error around the mean.*

Cumulative Distribution Frequency

Stereotype ER indexes computed across different time bins (see Figure 4) was analyzed in a factorial mixed ANOVA with Social Condition as between-factor and Bins (5) as within-factor.

Again, Social Conditions did not differ in their ERs (*F*<1). A marginal effect of Bins, *F* (4,492) =1.96, $p=.10$, $\eta^2=.02$, indicated (as seen in Figure 4) that the stereotype effect decreased across participants' RTs, suggesting that stereotype bias was lower during the slowest responses. Contrary to

what was expected, this Bin effect was not moderated by Social Condition; there was no interaction observed for Bin x Social Condition, F (4,492) =1.54, p=.19, η^2=.01.

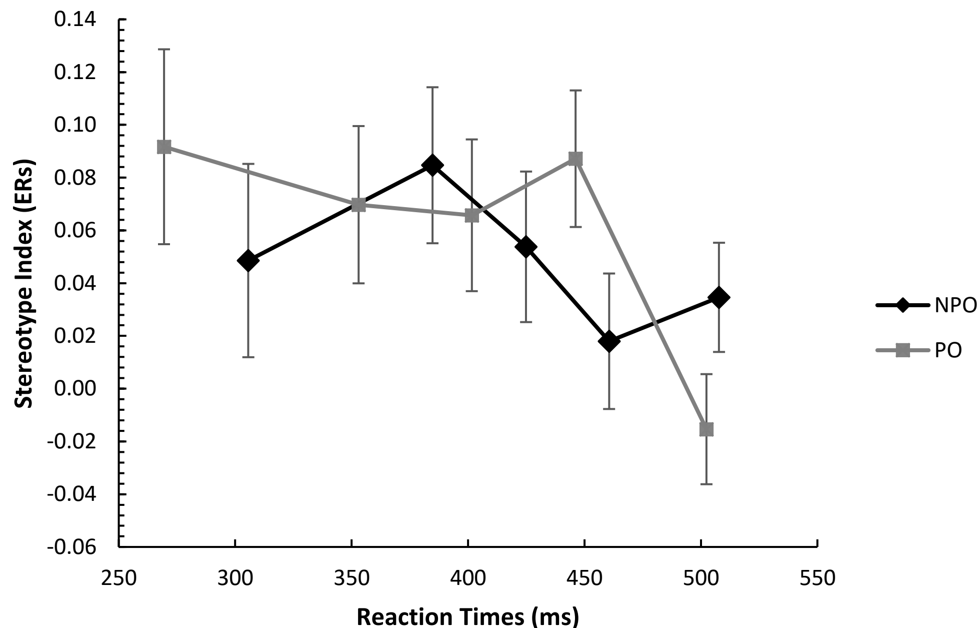

Figure 4. *CDF plots with Stereotype index (calculated based on Error Rates) for Social Condition (NPO vs PO) and Prime (Black vs White). Each marker represents a bin, which is associated with a Reaction Time Average. Error bars denote one standard error around the mean.*

Experiment 1- Discussion

Self-report data indicated that participants in PO felt more accompanied than participants in NPO, which corroborated our expectations for the SF manipulations. Additionally, the results of this analysis suggested that the manipulations did not impact how participants felt about being evaluated and observed. As such, it was unlikely that any evidence of SF effects was confounded with other effects, such as evaluation apprehension (Cottrell, 1972; Cottrell et al., 1968; Henchy & Glass, 1968; Belletier et al., 2015).

Participants in PO also reported feeling that the task required more cognitive capacity in comparison in those in NPO. There were significant negative correlations (r=-.21, p=.02 for *C Component* for Black primes, and r=-.20, p=.02 for *C Component* for White Primes) between this self-report rating scale and the control PDP components. This likely indicates that what was thought to be measuring capacity was actually measuring how depleted participants felt after executing the task.

As typically happens with the WIT, a stereotype effect was identified in the participants' responses (ERs). However, contrary to expectations, this effect was not moderated by social presence. Thus, these data did not replicate Lambert et al. (2003). A preliminary direct reading of this null result is that there is no SF effect in stereotyping and that what was previously found in Lambert et al. (2003) was accountability effects.

An alternative explanation is that the present manipulation (PO vs NPO) was subtler, and so the effects were less blatant. This possibility may have support in the tendency of those in PO condition to make more errors than those in NPO condition. Also, in accordance with Lambert et al. (2003), participants in PO seemed to have less control over their responses than those in NPO.

Evidence also suggests that PO may have induced different processing than NPO. Participants in the NPO condition showed an effect of Prime on the PDP *C Component*. This suggests that less control was exerted in the trials with Black primes than those with White primes (i.e., Black primes led to less discriminability between Guns and Tools). This effect was not previously documented in the literature (Payne, 2001; 2005; Lambert et al., 2003). The effect challenges the assumption of Payne et al. (2005) that Primes do not affect target perception in the WIT—a necessary assumption for believing that stereotype errors are dependent upon control failure (i.e., stereotype activation only affects responses when there is a lack of necessary attentional resources).

These data suggest that not only is control slightly higher in NPO than PO, but also the type of control that is activated in the NPO condition is affected by Primes. As such, participants in NPO and PO seemed to deal differently with the task. If this is a true effect in the present study, the data have likely documented an SF effect (although at different levels of analysis) that addresses the components of the responses and not the responses themselves (i.e., how responses are reached). When alone, the way participants exerted control over their responses depended on the Prime with which they were presented (i.e., they could more easily discriminate tools from guns when the prime was a white portrait than when it was a black portrait). This means that it was stereotype activation that was modulating the degree of correct target detection. This type of control, which depends on stereotype activation, is different from the type of control that was assumed to guide participants' responses in the PO condition.

Another result that seemed to emerge due to the NPO condition was the evidence of stereotyping in participants' RTs. That was not found in the PO condition nor in other studies in the literature. Payne (2001) suggests that stereotyping is only observed in RTs when participants have no

time restriction. In the NPO condition, even having responses restricted to 550ms caused participants to respond more slowly to the Black-Tool association than to any other association. Importantly, compared with participants in PO, those in NPO were slower overall in their correct responses, suggesting that their correct responses required more time to achieve than did those in PO. This is yet more evidence that SF effects occurred in a phase of processing that did not allow the detection of differences at a response level (i.e., accuracy level).

Together, the two differences in processing observed in the NPO condition (i.e., slower responses to the Black-Tool trials and less discrimination between tools and guns for Black than for White trials), may suggest that PO and NPO conditions differ not in the amount but in the type of control they exert over the task. This hypothesis is followed throughout this thesis.

In sum, the results did not show reliable SF effects on stereotype bias. Therefore, the results of Lambert et al. (2003), which detected higher stereotype bias in social contexts of higher-order, were not replicated, and nor were those of Castelli and Tomelleri (2008), which detected lower Stereotype bias in PO. However, it cannot be stated on the basis of the present data that PO does not interfere with stereotyping. Rather, it points to differences in the control exerted in PO versus NPO conditions. The PDP component suggests a tendency for less general control in PO, while in NPO control may be more specifically directed toward one set of trials than another (as it is dependent on the Prime). This indicates that Stereotype information is dealt with differently in different social conditions.

Before drawing further conclusions from this data, it was necessary to attempt a replication of the effect detected in IAT by Castelli and Tomelleri (2008).

Experiment 2 - Implicit Association Test with SF conditions

This experiment replicated the procedure used to identify SF effects in an IAT as described in Castelli and Tomelleri (2008). As in that study, the PO was manipulated by using co-action and NPO by having participants alone in a room without the experimenter. Following the original paper, the IAT index (*d600*) was used as a measure of stereotyping.

However, this experiment went further than mere replication by addressing the mechanisms underlying stereotype activation and control. The PDP components, ERs, RTs, and CDF plots were computed and analyzed. Given a replication of Castelli and Tomelleri (2008), the PDP components were expected to help determine whether the differences between PO and NPO when performing the IAT were also related to differences in control levels or to differences at the activation level. The

answer to this was expected to explain the opposing results seen for this task versus the WIT. Additionally, the CDF plots were expected to provide information about the kind of tradeoff between bias levels and RTs that sometimes occur. The moderation of this tradeoff by PO, furthermore, was expected to add to an understanding of how activation and control interact in this task, and how it is modulated by PO.

Taking a conservative perspective, and despite the results of Experiment 1, the results of Experiment 2 were hypothesized to match those obtained by Castelli and Tomelleri (2008): less stereotype bias in PO than in NPO. Nevertheless, following Experiment 1, this analysis focused on understanding whether individuals in different social conditions can differ not only in the level of stereotype bias exhibited but also/instead in the type of control they exert over their responses.

Experiment 2 - Methods

Participants and Design

A total of 119 non-black participants (10 men) with an average age of 20.97 (*SD*=5.89), enrolled at the ISPA (Lisbon, Portugal), volunteered to took part on the experiment. Social Condition (NPO vs. PO) was manipulated between participants and the IAT block (Incongruent Vs Congruent) within participants. The sample size was determined using the procedure employed by Westfall, Kenny, and Judd (2014): a minimum sample size for a power of .80, based on the main effect size of SF in Castelli and Tomelleri's (2008) Experiment 1 (*d*=.30) and the number of trials (120 critical trials).

Procedure

After participants gave their informed consent to participate in an experiment related to words and images evaluation, experimental sessions were scheduled in groups (PO) or individually (NPO) to manipulate Social Condition. In NPO, each participant did an IAT in a room without the experimenter. For the PO condition, participants arrived in groups of 6 to 10 and started the task at the same time in a room set up for the experiment. As in Experiment 1, the researcher was present, working on their own computer and positioned in such a way that it was impossible for them to monitor the participants' computer screens.

Each participant read the IAT instructions (presented on a computer screen), which explained that they were going to do a task that consisted of identifying words and images; they would evaluate words as Positive or Negative, and images (e.g., a Face) as Black or White.

The Race IAT comprised 5 blocks, as in Castelli and Tomelleri (2008). In the first block, participants evaluated 30 portraits (15 Black and 15 White) as Black (left key) or White (right key). In the second block, participants evaluated 30 positive words (e.g., hug, friend, love, kiss, child, embryo, fairy, fertile, flower, harp, oasis, paradise, tenderness, truth, virtue), using the left key to indicate Positive, and 30 negative words (e.g., infection, catastrophe, expulsion, ambush, degraded, corpse, wound, coffin, poison, bomb, rubble, avalanche, snot and maggot), using the right key to indicate Negative (see Figure 5).

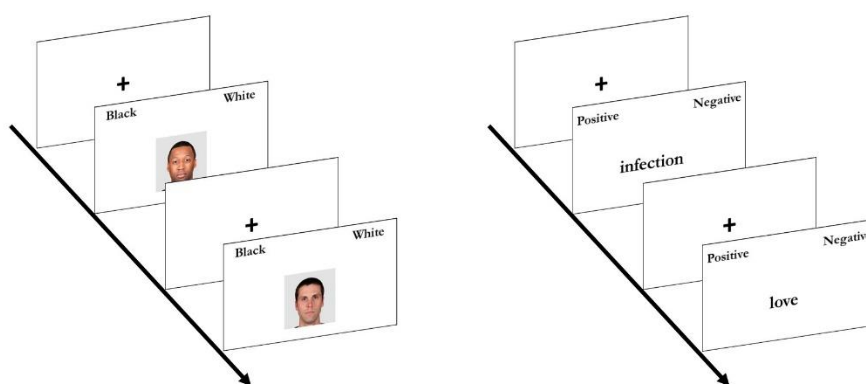

Figure 5. *IAT trials. The left side represents Block 1, in which participants evaluated portraits. The right side represents Block 2, in which participants evaluated words.*

In the third block, participants evaluated the same portraits and words as in blocks one and two, but with the categories Black and Positive presented on the same/left side of the screen, using the same response key, and the categories White and Negative on the same/right side, using the same response key. This block was Incongruent with the Black-Negative stereotype association. The fourth block was identical to the first block, but the Black and White categories switched positions; their respective keys were also switched. The fifth and last block was identical to the third block but this time the Black and Negative categories were presented on the same side (right key), while the White and Positive categories were presented on the other side (left key); this arrangement promoted congruency with the Black-Negative stereotype association (see Figure 6). Each IAT trial began with a fixation point of 500 ms followed by a target (portrait or a word) to which participants had to respond to proceed to the next trial. Note that contrary to the WIT in Experiment 1, the IAT did not have any time restrictions imposed on participants' responses.

After completing their IATs, participants received their credits and were dismissed.

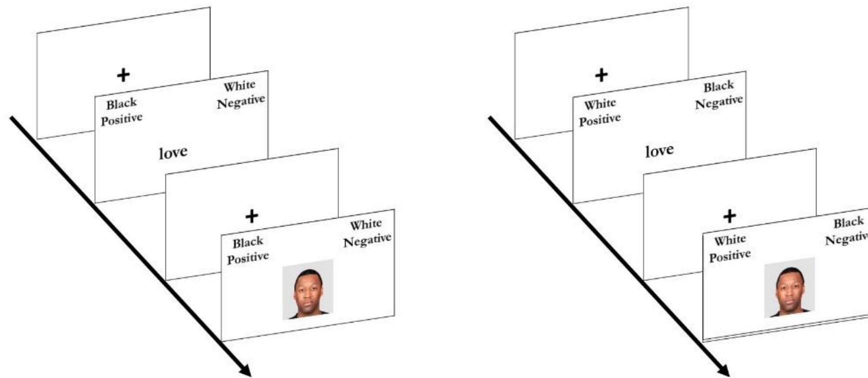

Figure 6. *IAT trials. The left side represents Block 3 (Incongruent). The right side represents Block 5 (Congruent).*

Dependent Variables

All dependent variables, detailed below, were extracted from the critical blocks (i.e., Incongruent (block 3) and Congruent (block 5) blocks).

D600 index is a standard index used in IAT studies. Its computation was achieved by following the steps outlined by Greenwald et al. (2003): 1) Eliminate trials with RTs greater than 10,000 ms and participants who had RTs slower than 300 ms in more than 10% of their trials; 2) Compute the mean of correct RTs for each critical block (i.e., Congruent and Incongruent) and the standard deviation (*SD*) of the two blocks together; 3) Replace each incorrect RT with the block mean and add 600 ms; 4) Compute the average of the critical blocks and the difference between those averages; 5) Divide the difference by the *SD* of the two blocks. By following these steps, the *d600* index for each participant was created; higher values indicate a more implicit and negative attitude towards Black people.

Error Rates (ERs) were calculated for each block (Incongruent and Congruent), where participants' errors were divided by the total number of recorded trials.

PDP Components were based on Stewart, von Hippel, and Radvansky (2009); PDP *C* and *A Components* were calculated for each participant. The PDP *C component* (probability of correctly evaluating the target) was calculated by subtracting the probability of correct responses in congruent trials by the probability of incorrect responses in incongruent trials. The PDP *A component* (probability of responding in a stereotypical way when control fails) was calculated by dividing the probability of incorrect responses in incongruent trials by (1-C). Because the percentage of incorrect responses was low, the Snodgrass and Corwin (1988) correction was applied so that PDPs could be used. This correction, used to cope with ceiling effects, simply adds a "half" correct trial (i.e., it adds .5 to the total number of correct responses), and one observation to the denominator (i.e., total number of responses). In this way, PDP components could be calculated even in tasks with low ERs.

Reaction Times of Correct Responses (RTs) was computed as an average of correct RTs for each block, which resulted in two RT averages (Congruent and Incongruent) for each participant.

CDF plots were used to analyze stereotype bias through RT distributions; the two correct-response (Incongruent and Congruent blocks) RT distributions were split into six bins for each participant. Then, the difference between Incongruent and Congruent bins for each specific bin (i.e., $Bin1_{Incongruent}$-$Bin1_{Congruent}$; … ;$Bin6_{Incongruent}$-$Bin6_{Congruent}$) was computed, where higher values indicated stronger bias.

Experiment 2 - Results

A preliminary analysis of ERs and RTs identified six participants with ERs above 90%, an abnormal value for a task without time restrictions. Thus, these participants were not included in the analysis, leaving 57 participants in NPO and 56 in PO.

d600

First, the *d600* index was compared between Social Conditions. Analysis revealed that participants in PO (M_{PO}= .27; SD=.35) had a higher *d600* index than participants in NPO (M_{NPO}= .13; SD=.30), $t(111)=-2.27$, $p=.03$, $d=.43$, (see Figure 7). This effect was in the completely opposite direction of that detected in Castelli and Tomelleri (2008). While Castelli and Tomelleri's (2008) results suggest a stronger implicit negative attitude towards black people (i.e., Stereotype Effect) in NPO, the present results suggest the opposite: a stronger implicit negative attitude towards black people in PO.

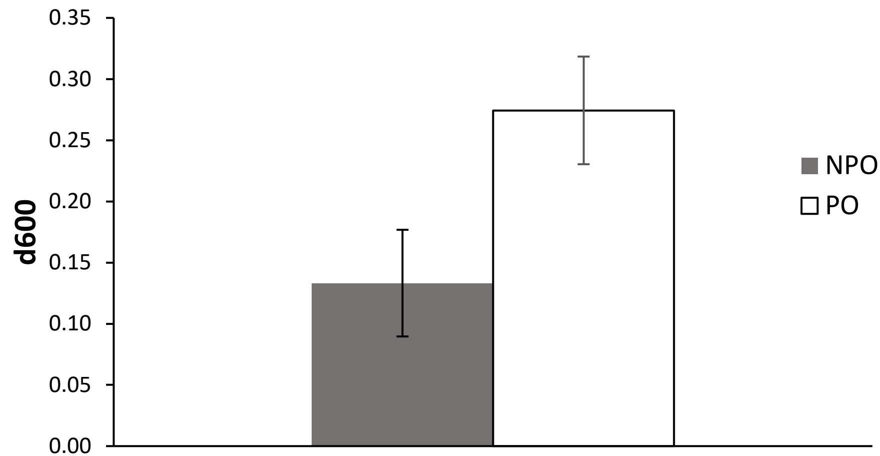

Figure 7. *d600 index in NPO and PO. Error bars denote one standard error around the mean.*

Error Rates (ERs)

ER analysis (see Figure 8) showed no effect of Block (*F*<1), of Social Condition (*F*<1), nor of Block x Social Condition interaction (*F*<1). This suggests that the stereotype effect in IATs is not captured in ERs.

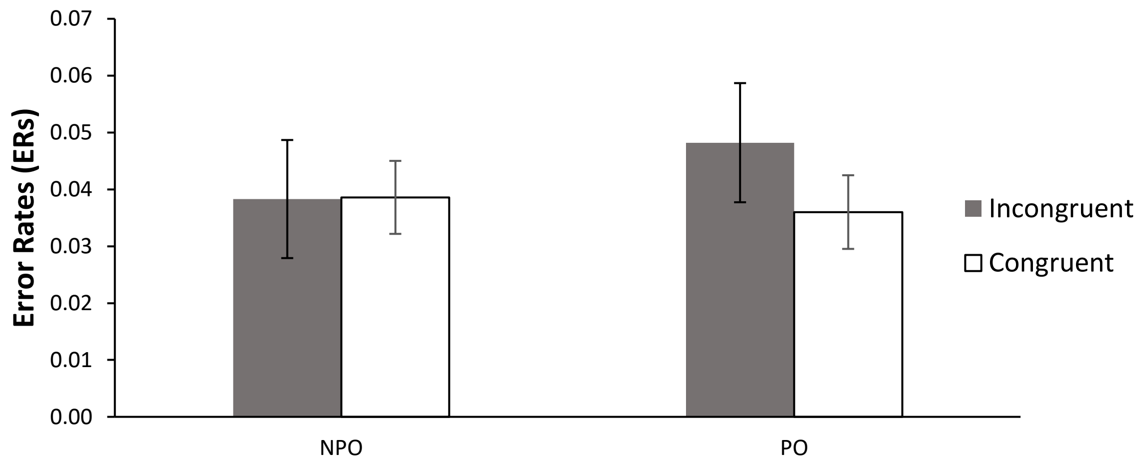

Figure 8. *Error Rates (ERs) for Blocks (Incongruent vs Congruent) in both NPO and PO. Error bars denote one standard error around the mean.*

PDP Components

Despite no effect being observed in ERs, a separate analysis was conducted for each PDP component in an attempt to detect differences in how Social Condition modulated the *C* and the *A Components* of participants' responses (see Table 3). Results indicated that Social Condition manipulation did not impact the *C Component* (t<1), nor the *A Component*, t(111)=-1.31, p=.78, d=.20. This suggests that the influence of control processes and stereotype activation on participants' responses were likely the same in both Social Conditions.

Table 3

PDP component estimates (Standard Deviation) for Social Condition

Component	**Social Condition**		**Overall**
	NPO	*PO*	
C	.91(.10)	.90(.10)	.90(10)
A	.49(.20)	.53(.21)	.51(.20)

Reaction Times of Correct Responses (RTs)

RTs analysis seem reductant given the effects that occurred with *d600* (index is strongly based on RTs). However, the *d600* also considered errors, and so different results can show how clearly the effect can be detected by only one component of participants' responses (RTs vs ERs). Thus, RTs were analyzed using a mixed ANOVA with Block (Incongruent vs Congruent) as within-factor and Social Condition (NPO vs PO) as between-factor (see Figure 9). A main effect of Block, F (1,111) =16.24, p<.001, η^2=.13, showed that participants were slower in the Incongruent block, M=829.71; SD=355.76), than in the Congruent block (M=722.69; SD=151.01) (i.e., the stereotype bias was detected in RTs). More importantly, this effect was marginally moderated by Social Condition, F (1,111) =2.90, p=.09, η^2=.03, suggesting that Stereotype bias is prone to be higher in PO.

Additionally, no reliable main effect of Social Condition, F (1,111) =2.70, p=.10, η^2=.02, was detected, suggesting that participants in PO (M=812.39; SD=231.33) may be slower to respond than participants in NPO (M=740.02; SD=231.43).

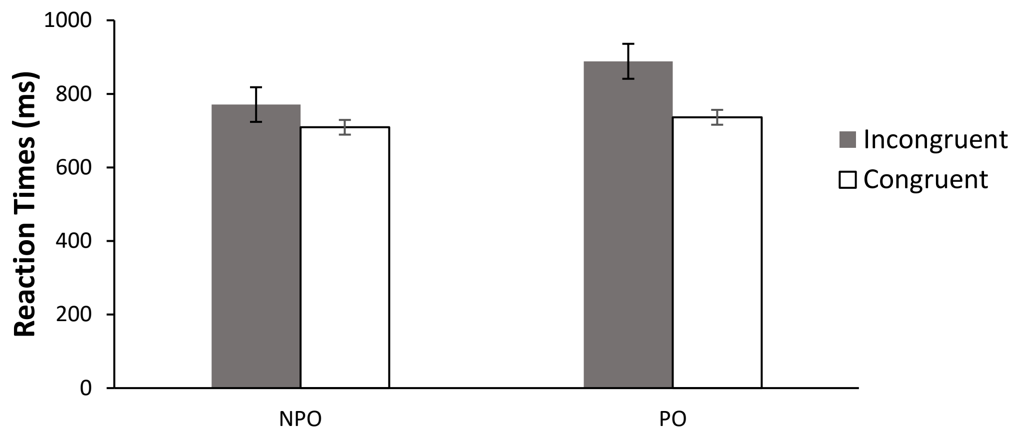

Figure 9. *Reactions times (RTs) for Blocks (Incongruent vs Congruent) in both NPO and PO. Error bars denote one standard error around the mean.*

Cumulative Distribution Frequency (CDF)

CDF plots for Incongruent-Congruent differences (see Figure 10) were analyzed using a mixed ANOVA with Social Condition as between-factor and Bins (6) as within-factor. The ANOVA detected a main effect of Bins, F (5,555) =14.26, p<.001, η^2=.11, indicating that the level of Stereotype bias (like other types of bias effects) increased over time (i.e., the bias was stronger when individuals delayed their responses). However, this relation was moderated by Social Condition, F (5,555) =3.07, p<.01, η^2=.03, (see Figure 10), suggesting that control was more efficiently implemented in one condition than in the other. Results indicated that participants in the NPO condition showed less evidence of interference, even in their delayed responses; it is likely that they were more efficiently controlling stereotype interference during their task. As Figure 10 shows, the stereotyping effect increased across RT distribution more sharply in PO (higher slope) than in NPO.

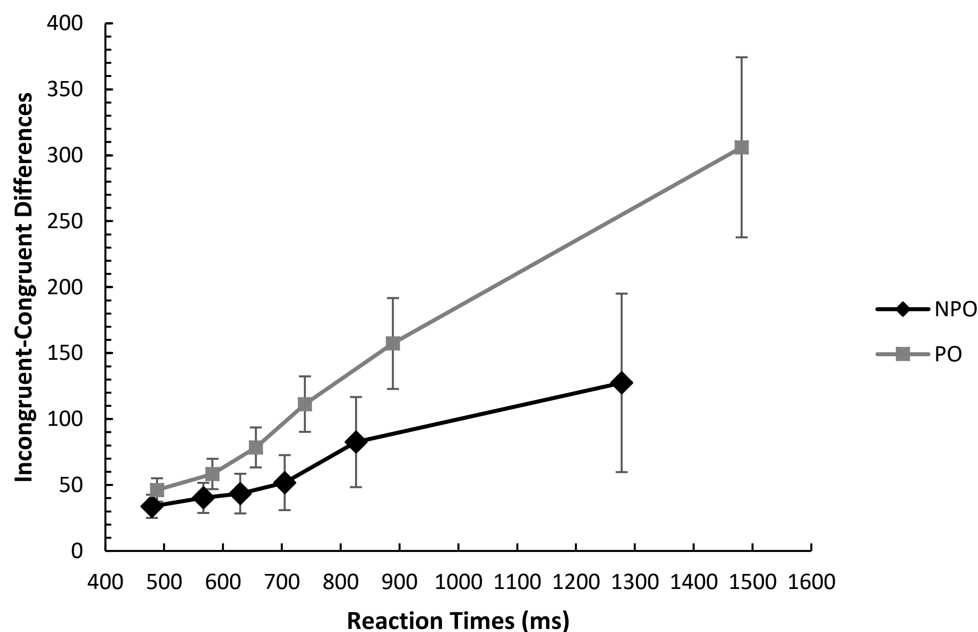

Figure 10. *CDF Plots with Incongruent-Congruent RT differences for Social Condition (NPO vs PO). Each marker represents a bin that is associated with a Reaction Time Average. Error bars denote one standard error around the mean.*

Experiment 2 - Discussion

On the basis of Castelli and Tomelleri's (2008) results, it was expected that more evidence of stereotype bias would be observed in NPO. However, the present replication of their Experiment 1 revealed the opposite pattern: the SF effect indeed occurred but showed more stereotype bias in PO than in NPO. This suggests that the conclusions reached by Lambert et al. (2003) regarding the WIT were replicated in the present study with the IAT; evidence of SF was shown because those in alone condition are less susceptible to stereotype effects than those in the presence of others.

The question of why the original study results were not replicated can be answered by addressing the differences between that study and the present one. First, there was a cultural difference (Portuguese vs Italian). It is possible that the social norms of equality are more accessible in Italian culture than in Portuguese culture. Thus, the results of Castelli and Tomelleri (2008) may not be showing only SF effects but also specific accessibility to norms in that particular sample.

Second, there was a difference in how social conditions were operationalized. While their paper (Castelli & Tomelleri, 2008) suggested that they used a true alone condition, the authors revealed in personal communication that a researcher was present in their NPO condition. This discrepancy

between their NPO condition and that of the present study is relevant: the presence of a researcher may have produced an evaluative apprehension condition. This may help explain the difference in means of the stereotype interference observed in the studies, namely in the reported *d600* averages. While the averages in PO (.27 versus .29) were similar, there was a much greater discrepancy between the averages in NPO: .13 (present study) versus .46 (Castelli and Tomelleri, 2008). This simple difference may be enough to change the results, as stated in Chapter I and also described by Markus (1978). Furthermore, in their meta-analysis of 241 SF studies (289 different experiments), Bond and Titus (1983) identified 131 experiments that reported an NPO condition with a researcher present, and 41 experiments with uncertain or unclear NPO parameters (i.e., impossible to determine whether a researcher was present).

Although the PDP approach was used as a means to understand whether the stronger stereotype bias in PO was caused by more stereotype activation or by less control, this measure quickly revealed a problem. Because PDP is based on accuracy data, and the effect in IATs is measured by *d600*, which relies more on RTs (see Klauer & Voss, 2008; Ito et al., 2015), the PDP measure was not sensitive enough to allow the reliable detection of differences in PDP components.

To understand the present data, it is important to note that SF is more clearly identified with the *d600* than with a mere comparison of RTs. Typically, stereotype bias in IATs is likely to be detected in RTs, as accuracy data are subject to less interference. However, that was not the case in these data; direct analysis of RTs detected only a marginal effect. This suggests that because the *d600* measure penalizes errors in its computation, it detected the SF effect more easily (i.e., it is easier to see how PO modulates the processes underlying stereotyping when RTs and response accuracy are put together). The relevance of this will become more apparent as the SF effect is further explored in the next chapter.

The CDF analysis of Stereotype bias (measured in RTs) across RT distribution was also conducted to explore the ways in which time can inform us about how SF modulates stereotyping. This analysis revealed that bias strengthened with an increase in RT (i.e., the stereotype bias in IATs is stronger in slower responses than in faster responses). This pattern has also occurred in other interference data, such as that for Stroop tasks (Bub, Masson, & Lalonde, 2006; Sharma et al., 2010) and flanker tasks (Ridderinkhof, Wildenberg, Wijnen, & Burle, 2004; Ridderinkhof, Scheres, Oosterlaan, & Sergeant, 2005; Davranche & McMorris, 2009). As previously detected by Sharma et al. (2010), a change in this pattern suggests that one condition is using the time for controlling

interference more efficiently than another condition. However, the pattern of results in the present study was the opposite to that observed in Sharma et al. (2010): Stereotype bias increased across RTs more in PO than in NPO, which caused participants in PO to exhibit more stereotype bias in the later bins relative to participants in NPO. This occurred either because stereotype activation was stronger in PO or because levels of control were higher in the NPO conditions (Wagstaff et al., 2008).

At some level, Experiment 2 offers information in the same direction as Experiment 1. In both studies, less control in PO was observed, which provides clear evidence of stereotyping in Experiment 2 but not in Experiment 1. Furthermore, the data from both studies show that to understand SF effects, we must also attend to their components (RTs and Accuracy) and, more importantly, to the dynamic features of how those components interact with one another. The two experimental tasks (WIT and IAT) differ in the variables that can explain the interference in stereotype activation; whereas stereotype bias in the WIT (with time restriction) is explained through response accuracy, stereotype bias in the IAT (without time restriction) is explained by RTs. It is thus possible that the dynamic between time and accuracy is differently detected in each of these tasks. The conjunctions of RT and ER effects in the d600 in IATs, suggest that indexes that combine the two are more likely providing information about SF effects in stereotyping. In the next chapter, a new way of approaching these data is offered.

Chapter IV
A Diffusion Model approach

In the previous chapter, we stated that SF effects on stereotyping may not be easily detected when the participant's ERs and RTs are analyzed in separate. The results of experiment 2 (IAT) offered a good illustration of this assumption, since we did not find SF effects on stereotyping in isolated analyses for ERs and RTs. However, using the *d600* index, which combines both variables, we found more stereotyping in PO than in NPO.

To systematically address this question, we suggest the use of Ratcliff's (1978) Diffusion Model (DM), which combines RT and ER data in the computation of several indices/parameters. In this chapter, after presenting the DM rationale and requirements, we will explain how parameters are calculated and interpreted in different tasks. Moreover, we will also present examples of previous studies using stereotype tasks, in which data was analyzed using DM.

Finally, to further explore how the DM can shed new light on the intervening mechanisms of stereotype tasks, we will reanalyze Experiment 1 (WIT) using this method. The application of DM to the data of Experiment 2 was not possible due to the lack of sufficient ER variability. To address the pending questions of Experiment 2, we performed Experiment 3 using the same IAT task, but with more trials and a response time restriction in order to meet the DM data requirements. Therefore, in experiment 3, we aim to replicate the results of Experiment 2 and reanalyze the data using the DM.

The Diffusion Model-Computation of mechanisms through responses and reactions times.

In the literature, we can find evidence that stereotype effects are not only detected on responses (accuracy data), but can also be found in RT analysis (e.g., Fazio et al., 1995). Compared to these more classical analyses, the DM (Ratcliff, 1978) presents a clear advantage by allowing the estimation of different parameters, associated with psychological mechanisms, based on both responses' accuracy and RTs. This approach also contrasts with other estimation models such as PDP (Jabocy, 1991) and Signal Detection Theory (Swets, 1964), which are only based on accuracy data, neglecting RTs data.

The DM has been successfully applied on a variety of binary decision contexts (e.g., Voss, Rothermund, & Voss, 2004; Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007; Voss, Rothermund, Gast, & Wentura, 2013), including those specifically related to stereotyping (Correl et al., 2015; Pleskac, Cesario, & Johnson, 2017). For example, Correll et al. (2015) and Pleskac et al. (2017), used DM estimates to examine perceptual/interpretational processes in the First-Person Shooter Task (FPST). This task is similar to the WIT, but the participants instead of having to decide if a Gun/Tool was presented, they have to decide to Shoot/Not Shoot armed/unarmed individuals.

To understand how the DM could provide insights about SF effects and stereotyping mechanisms, it is first necessary to understand how this complex model generates its parameters. Within the DM, accuracy data is used to distinguish Correct and Incorrect decisions. These decisions are represented in Figure 11 as the Correct Decision Threshold (upper threshold) and the Incorrect Decision Threshold (lower threshold). A specific RT distribution is then associated with each response threshold (e.g., decision), as it can also be seen in Figure 11. The DM assumes that decisions are made by capturing information/evidence (continuous sampling) until one of the thresholds is reached. Based on this assumption, using the response ratios and its temporal distribution, we can estimate how fast a decision threshold is reached (*v Drift-Rate*), how much evidence is required to reach the thresholds (*a Threshold Separation*), and which threshold needed less evidence to be reached (*zr Relative Starting-Point*).

In Figure 11, the *Drift-Rate* (*v*) is represented by an arrow (or slope of the sampling process) that is going in the direction of the Correct Decision Threshold. The *Drift-Rate* direction and slope, directly depends of the response ratios and RT distributions, such that faster RTs and a higher number of Upper threshold decisions (in this example correct responses) contributes to higher and positive *Drift-Rates values*. Therefore, the *Drift-Rate* (*v*) represents the average of information/evidence that was extracted from a stimulus for each unit of time in order to make the decision (Johnson, Hopwood, Cesario, & Pleskac, 2017). In other words, how easy it was to reach a decision in specific conditions/trials. Therefore, we can expect congruency effects on *Drift-Rates* (Schmiedek, Oberauer, Wilhelm, Süß, & Wittmann, 2007; Voss, Rothermund, Gast & Wentura, 2013), since incongruent trials have information that can be used as supporting evidence for two different decisions (making the decision more difficult), whereas the information in congruent trials points to the same decision (facilitating the decision).

Differently of the *C Component* in the PDP (Jacoby, 1991) and the *d'* parameter in Signal Detection Theory (Swets, 1964), *Drift-Rate (v)* is more than a measure of performance. By integrating RT data (Klauer & Voss, 2008), *Drift-Rate* can also be considered a measure of efficiency, for it provides information on how good was the performance of the participant and how fast that level of performance was reached. For example, Schubert, Frisckorn, Hagemann and Voss (2016) showed in multiple tasks that more intelligent participants had higher *Drift-Rates*, suggesting that the parameter capture the participants' efficiency.

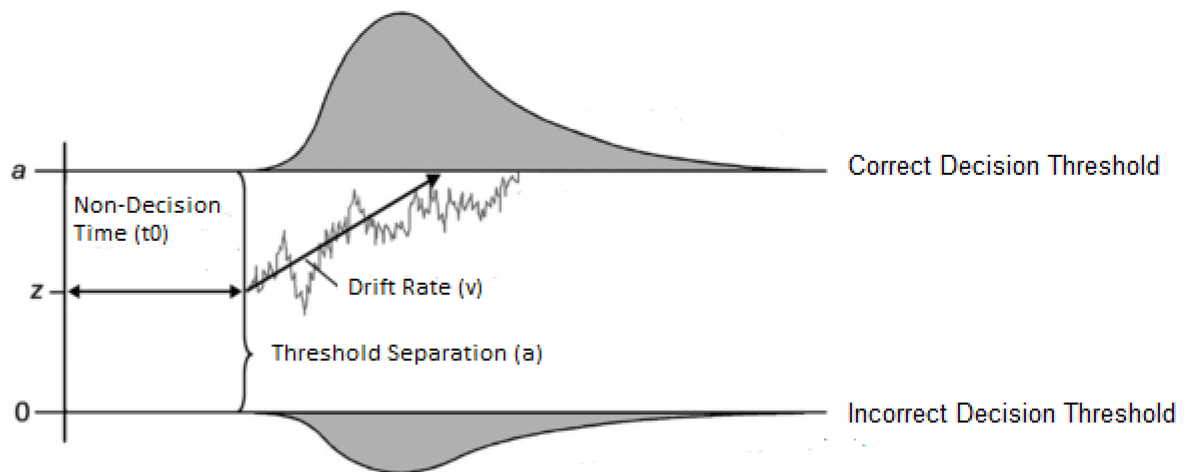

Figure 11. *Diffusion Model illustration*

Threshold Separation (*a*), as the name indicates, represents the distance between the two decision thresholds. This distance can be interpreted as the amount of information that was considered for a decision and is directly related with the time participants spend evaluating the stimulus (Voss, Voss & Lerche, 2015; Johnson, Hopwood, Cesario, & Pleskac, 2017). Since the bottom threshold is usually represented by 0 and the upper threshold by a (see Figure 11), *a* also refers to the distance between the two decision thresholds. A higher *Threshold Separation* signals that more information was used during the decisional process and since a higher amount of information leads to better decisions, higher values for threshold separation are usually associated with more correct responses. Moreover, a higher *Threshold Separation* indicates a cautious decisional process, that required more time to collect sufficient evidence. For instance, Voss, Rothermund, Gast and Wentura (2013) showed that when

participants were instructed to be more cautious in specific trials, a higher *Threshold Separation* was detected for those trials. Furthermore, when participants were instructed to prioritize speed (e.g., give faster responses), they presented a lower *Threshold Separation*, than the participants that received instructions to prioritize accuracy (e.g., respond accurately) (Ratcliff & Rouder, 1998; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008).

Since *Drift-Rates* (*v*) and *Threshold Separation* (*a)* establish a relation between response (e.g., accuracy) and RTs, it is important to distinguish the two parameters for they are associated with different concepts. Higher *Drift-Rates* are associated with faster and more correct responses. Therefore, *Drift-Rate* is a measure of easiness/efficiency to achieve a decision. This allows the experimenter to use this parameter to contrast between different decisions (e.g., Gun/Tool). Moreover, *Drift-rates* can also be used to compare how the decisions are achieved for congruent (stereotypical information) and incongruent (counterstereotypical information) conditions. Based on the assumption that the decision is easier in congruent trials, higher *Drift-Rates* are expected than in the incongruent trials. On the other hand, a higher *Threshold Separation* is associated with slower and more correct responses. *Threshold Separation* can then be used as a measure of caution during the decisional process. It is important to note that *Threshold Separation* is a single parameter that results from the combination of the information used in both decision alternatives, i.e., the quantity of information required to reach the decisions and not a decision in a specific direction. Therefore, *Threshold Separation* is more related to decisional styles or strategies (impulsive vs cautious) and, unlike the *Drift-Rate* parameter, it does not allow comparisons between the decisions themselves (e.g. Gun/Tool). For instance, higher *Threshold Separation* for Black trials than for White trials, when the decisions are between Gun and Tool (e.g., WIT task), indicate that participants were more cautious in Black trials, which can indicate a possible tactic in order to avoid prejudice responses (see Pleskac, Cesário & Johnson, 2017; which will be presented further). This information can be relevant for our Experiment 1, since NPO participants showed stereotype effects on ERs and on RTs, suggesting an attempt to be more cautious on counterstereotype trials.

The DM parameter *Starting-Point* (*z*), as seen in Figure 11, refers to the point between the two decision thresholds from which the *Drift-Rate* begins. For example, in Figure 11, the *Starting-Point* is displayed is in middle of the two decision thresholds, indicating that at the beginning of the decisional processes, both decisions have the same probability to be made. Because participants vary in terms of *Threshold Separation* (distance between decision thresholds) the *Starting-Point* cannot be directly compared between participants. To solve this issue, the *Relative Starting-Point* (*zr*) must be computed

by dividing the z value by the threshold separation (z/a). The values for the *Relative Starting-Point* (zr) vary between 0 and 1, with values higher than .5 indicating a tendency to the decision threshold a (e.g., upper threshold) and values lower than .5, indicating a tendency to the lower threshold decision (e.g., lower threshold). Importantly, these values represent a tendency towards one of the decision thresholds in the absence of evidence (before the decisional process starts).

The *Relative Starting-Point* of the DM can be interpreted as a general prior decision bias, or an individual tendency to respond in one direction versus another. For instance, the literature shows that the *Relative Starting-Point* (zr) can be influenced by motivational changes, such as the presence of a reward; e.g., trials with higher rewards lead participants to have a higher relative starting point to the correct response, than trials with lower rewards (Voss, et al., 2004). Moreover, if the decision thresholds are set to Shoot/No Shoot decisions, instead of Correct/Incorrect decisions, this parameter works as the *A Component* of the PDP (Jacoby, 1991) or the c parameter of the Signal Detection Theory (Swets, 1964), giving us access to a measure participants' bias to Shoot/No Shoot.

Finally, the DM estimates the time outside of the sampling of information and decision processes. This parameter is called *Nondecision Time* ($t0$) and it incorporates preparatory decisional processes (e.g., stimulus-encoded features involved, switching costs/activation of the correct task set, directing of the attention, configuration of the working memory) and post-decisional processes (e.g., translation of the decision into a motor action). In the literature, an increase in $t0$ values has been claimed to capture an increase in response execution interference (Voss, Rothermund, Gast & Wentura, 2013), and to be related with switching costs (costs related with changes of task goals between trials, Schmitz & Voss, 2012), and also as a measure of time execution of the response (asking participants to use only one finger for both response keys leads to a higher $t0$; Voss et al., 2004). Since this parameter results from a combination of multiple processes, *Nondecisional Times* should only be interpreted when experimental conditions are set to sufficiently isolate the target process. For example, an experiment by Schmitz & Voss, (2012) manipulated trial order, so that the differences in nondecision times could have only been explained by switching costs.

It is extremely important to note that, since the DM requires a valid RT distribution for each of the decision thresholds (e.g., Correct and Incorrect decisions/ Gun and Tool decisions), a high number of trials is required for the correct estimation of the parameters. The DM also requires response variability to ensure that the parameter estimates are not biased (Lerche, Voss & Nagler, 2017). These requirements invalidate the use of the DM on any task in which the effects are found on

just one specific metric. For instance, LDTs in which the effects are exclusive to RTs and not found on response variation (errors).

Although the DM was not specifically designed to accommodate a dualistic perspective, it is possible to establish parallels between DM parameters and the PDP (Jacoby, 1991). Klauer and Voss (2008) suggests that the *Relative Starting-Point* has the same rationale as the *A Component* of the PDP, for it measures an initial, automatic bias. Since the PDP assumes that control processes lead to performance, the same parallel can be made between the *C Component* of the PDP and the DM *Drift-Rate*. However, while the *C Component* is a measure of performance that relies on response accuracy, *Drift-Rate* is a measure of performance efficiency, for it combines accuracy and time. Note also that the *C Component* of the PDP, in WIT tasks, works as a measure of discriminability between two objects (Tool-Gun), while in the DM, *Drift-Rates* can be calculated for each object (i.e., decision; Pleskac, Cesario & Johnson, 2017). Moreover, since the *Threshold Separation* parameter of the DM functions as a measure of how much information/evidence is needed for decide, it could also be used as a measure of control that relies on activated information and time. *Threshold Separation* could then be indexing a type of control similar to the inhibitory control that was described by Amodio et al. (2004), according to whom, control depends on activation of information and time for inhibition.

In conclusion, the DM can be a useful tool to disambiguate the underlying processes of effects reported in the literature. For example, the DM analysis in Racliff, Thapar and McKoon (2004) showed that, in recognition tasks, the differences in performance between older and younger participants resulted from the speed of response execution (*Nondecion Time*) and not from cognitive processing differences (*Drift-Rate*). Moreover, as Voss, Nagler and Lerche (2013) suggest, some effects can go undetected in these separate analyses of RT and Accuracy, because they are scattered across the two measures. Therefore, these effects can only be detected using data analysis methods that combine both measures. This opens the possibility of new findings by applying the DM in stereotype tasks like Experiment 1 (WIT), in which it seemed that NPO and PO participants were affected by stereotype bias in different measures.

Diffusion Model on Stereotyping

Probably due to the complexity of the DM, we have only found two papers using this analysis with data from a stereotype task, namely the First-Person Shooter Task (FPST; Correll et al., 2015; Pleskac et al., 2017). The FPST is a task where participants see in each trial a background with an individual (Black or White) holding an object (Gun or Neutral object). The FPST simulates a threat

situation, in which participants must decide to shoot individuals holding guns and not shoot unarmed individuals. At a first glance, this task may look similar to the WIT, since it detects errors driven by the stereotype activation, leading individuals to shoot more unarmed black individuals, than unarmed white individuals. However, the WIT (contrary to the FPST) is a sequential priming task where memory plays an important role, since the target is presented on screen for only 100 ms. Moreover, since the FPST has a SOA[1](Stimulos Onset Asynchrony) of zero, the decision is made based on the current perception of an image.

Although there are substantial differences between the FPST and WIT, the studies that analyzed FPST with the DM, may help us interpret the DM parameters in our WIT experiments. We will now review the studies of the two papers previously mentioned.

Correl, Crawford, Wittenbrink and Sadler (2015)

Data from three experiments using the FPST were analyzed with the DM and 3 different parameters were estimated: *Drift-Rates* (*v*), *Relative Starting-Points* (*zr*) and *Nondecision Time* (*t0*). *Threshold Separation* (*a*) was not analyzed.

Since the authors considered that it was the tendency to Shoot or No Shoot that defines the impact of the racial Stereotypes (e.g., more shoot decisions for Black individuals), these decisions were set as the decision thresholds (instead of using Correct and Incorrect as thresholds). This way, the DM defined four different *Drift-Rates* (*v*), representing Shoot armed Black, Shoot armed White, Not Shoot unarmed Black, Not Shoot unarmed White. The comparison of *Drift-Rates* should capture the influence of stereotypes on the individuals' decision, allowing to determine how much race (Black vs White) and object (armed vs unarmed) helped the decisions to Shoot/No Shoot, (i.e., discriminating a dangerous situation in which participants should shoot, from a safe situation in which participants have no reason to shoot). If stereotypes affect the visual interpretation of the target for the Shoot/No Shoot decisions, we should expect higher *Drift-Rates* values for stereotype congruent decisions (Shoot an armed Black, don't Shoot an unarmed White), than for incongruent/counterstereotype decisions (don't Shoot an unarmed Black, Shoot an armed White).

The authors also computed two *Relative Starting-Points* (*zr*), one for Black and another for White trials. Each *Relative Starting-Point* represents the tendency to Shoot (values higher than .5) or Not Shoot

[1] Stimulos Onset Asynchrony is measure of the amount of time between two stimuli. While the FPST present the object and individual in at same time (SOA of 0). The WIT is sequential present a prime 100 ms before the presentation of the target (Object). Thereby have a SOA of 100 ms.

(values lower than .5), when the target is a Black or a White person. The comparison between Black *zr* and White *zr* should provide information on a possible prior bias for individuals to shoot a Black target in comparison with a White target.

The *Nondecision Time* parameter was also calculated for each of the four conditions (Shoot armed Black, Shoot armed White, Not Shoot unarmed Black, Not Shoot unarmed White). The detection of effects on this parameter allows to understand if the stereotype bias is also explained by motor response competition (*Nondecision Times*), or if it is only explained by cognitive mechanisms (*Drift-Rates).*

The data analysis of experiment 1, showed that *Drift-Rates* were higher for the trials with armed individuals, than for the trials with unarmed individuals, indicating that it was easier to make a Shoot decision, than a Not Shoot decision. As expected, the authors also report an interaction between Object x Race, indicating that the decision to Shoot was easier for armed Blacks, than for armed Whites, and that a decision to Not Shoot an unarmed white target was marginally easier than for unarmed Black. Together, these results show how stereotypes influence the way that participants disambiguate the visual information they are receiving. In the same experiment, *Relative Starting-Point* was not influenced by Object or Race, suggesting that there was no prior bias to decide to Shoot any of the targets. In addition, the *Nondecision Time* parameter was faster for armed individuals than for unarmed individuals, but no effect related to the race was detected, which suggests that stereotype interference only affected decisional processes and not response execution.

In experiment 2, authors consolidate the interpretation of the DM parameters with eye-tracking data. They showed that *Drift-Rates* effects were related to how quick participants stopped searching for more information to support their responses. This search was quicker on stereotypical trials, than on counterstereotypical trials. Data from this experiment, contrary to experiment 1, also showed an effect of Race on the *Nondecision Times* parameter, indicating that participants were faster to react to Black targets, compared to White targets.

In experiment 3, following Payne et al.'s (2005) procedure, participants were offered a second chance to respond. As reported in Chapter II, this procedure allows to discriminate between a race effect rising from an immediate influence over object perception (leading participants to give the same responses in their second attempt), from a race effect due to the failure of controlled processes (leading participants to change/correct their response in a second attempt). In this experiment, the authors showed that the differences found on the *Drift-Rates* computed for the first attempt, were associated

with the participants' responses (decisions) in the second attempt. This suggests that the stereotype effect found in the FPST was due to modifications on the participants perception of the object, leading them to make the same decision even when given a chance to correct it (which contradicts Payne et al.'s 2005 WIT results).

Pleskac, Cesario and Johnson (2017)

In this set of FPST experiments, Pleskac et al. (2017) replicated the DM analysis of Correl et al. (2015). The decision thresholds were set to Shoot and Not Shoot. The DM estimated 4 *Drift-Rates* (armed Black, armed White, unarmed Black and unarmed White trials), 2 *Relative Starting-Points* (Black and White trials) and 4 *Nondecision Times* (armed Black, armed White, unarmed Black and unarmed White trials). However, contrary to Correl et al. (2015), these authors computed and analyzed 2 *Threshold Separations* (Black and White trials), which were interpreted as response caution. Differences between Black-White *Threshold Separations* would mean that participants were more cautious (searched for more evidence) on specific type of trial (Black/White).

The results replicated the experiment 1 of Correl et al. (2015), but with a response time restriction of 850 ms. The *Drift-Rates* analysis detected the interaction Race (Black vs White) x Object (armed vs unarmed) indicating that it was easier to Shoot on trials with armed Black individuals, than on trials with armed White individuals. The Object main effect also indicated that it was easier to decide Shoot in the presence of an armed target, than to decide Not Shoot in the presence of an unarmed target. In conclusion, the *Drift-Rates* analysis seemed to capture the impact of stereotypes in the participants' decisions, so that a Black individual constituted extra evidence for Shoot decisions. The same experiment also reported differences in the *Relative Starting-Points*, showing that participants had a tendency to Tot Shoot on Black trials, compared to White trials. This suggests that participants were trying to avoid prejudiced responses.

The analysis of the *Threshold Separation* parameter revealed differences promoted by Race. This parameter had higher values for Black trials, than for White trials. This suggests again that participants were trying to not give prejudiced responses at Black trials, since they were searching for more evidence before their shooting decisions, i.e., spending more time in their decisions, when the trial has a Black individual.

A main effect for Object on *Nondecision Times* (quicker for armed trials) and an interaction Race x Object was also found, but not subsequently replicated in the other set of experiments.

In experiment 2, in order to decrease control, the response time was reduced to 630 ms. Moreover, the authors also added a scenario manipulation: Dangerous (Black/White man would be presented in a dangerous neighborhood) vs Neutral (Black/White man would be presented in a neutral scenario).

The results of experiment 2 also revealed a Race x Object interaction for *Drift-Rates*, showing that stereotypical information had a predominant role in how participants discriminated their decisions. However, this effect was not moderated by the scenarios (dangerous vs neutral).

The fact that in this experiment no differences promoted by Race (Black and White) were detected for *Relative Starting-Point* (contrary to experiment 1), suggests that the response time restriction (630 ms), prevented participants of using a prior bias in order to avoid shoot decisions in Black trials.

Also due to the time restriction, the values for *Threshold Separation* were also inferior relative to those found in experiment 1. This suggests that the parameter is capturing the opportunity of participants to individuate their response by searching for more evidence/information. However, because the time restriction was lower in this experiment, the opportunity to use this mechanism was also diminished. Despite the lower threshold separation, the race effect on this parameter remained significant, showing that participants were more cautious on Black trials than on White trials. Finally, the *Nondecision Times* was faster for Shoot than for No Shoot decisions.

In experiment 3, the authors set the response time restriction to 750 ms, hoping to find values for threshold separation between those reported on Experiment 1 (response time restriction of 850 ms) and Experiment 2 (response time restriction of 630 ms). In this experiment, half of the objects were blurred to make them more ambiguous (allowing to detect a need for more information). The danger of the scenario was now manipulated within-subjects, and not between-subjects.

The analysis of the *Drift-Rates* detected a stereotype effect; however, the parameter values were reduced by the blurred objects. This finding corroborates that *Drift-Rates* can be interpreted as a discriminability measure, displaying lower values, when it is more difficult to extract information from objects.

As expected, the *Threshold Separation* values (that again were higher at Black trials than at White trials) were between those detected in experiment 1 and 2. Indicating that this parameter was

influenced by response time restrictions and captured a need for information that could reflect a control mechanism that relies on the available time/information.

The *Nondecision Times* were again larger for Not Shoot decisions, than for Shoot decisions. Moreover, a scenario effect was now detected (unlike in experiment 2), with the dangerous scenarios leading to higher *Nondecision Times,* than the neutral scenarios. No effects were detected in the *Relative Starting-Points.*

Experiment 4 replicated experiment 3, but with a response time restriction of 630ms. More participants and trials were also added. *Drift-Rates* once more showed the stereotype effect. No effect for *Relative Starting-Point* was detected, possibly indicating no prior bias caused by Race. The fact that the effect of Race on *Threshold Separation* disappears, indicates that this parameter is sensitive to the restriction of response time. This further corroborates the interpretation that *Threshold Separation* is capturing a control mechanism that aims to avoid prejudiced responses. As before, the *Nondecision Times* were smaller for Shoot decisions than for No Shoot decisions.

In order to isolate the stereotype effects on DM from other possible effects, the author merged all their data (experiment 1, 2 , 3 and 4) and recomputed the DM parameters, without using data from conditions in which other manipulations were performed (i.e., ignoring conditions with blurred objects and/or the manipulation of scenarios) and concluded that: a) the *Relative Starting-Point* parameter showed a general prior bias to Shoot, that was not qualified by Race; b) the *Threshold Parameter* showed a direct effect of race. Participants were more cautious on Black trials, probably because they were correcting the influence of the stereotype. This interpretation of the parameter was corroborated by the fact that it was sensitive to response time restriction manipulations, so that higher response time restrictions lead to lower values of *Threshold Separation*, and consequently a lower capacity to avoid stereotype bias; c) the *Drift-Rates* showed that stereotype bias was represented by the interaction between Race and Object. In this interaction, higher values were observed for Shooting armed Black targets, than for White armed targets. Moreover, higher values were observed for the decision of Not Shooting unarmed White targets, compared to unarmed Black targets. This evidence suggests that the *Drift-Rate* parameter indicates how stereotype information can be used to quickly disambiguate the information/evidence for the participants' make their decisions; d) the *Nondecision Time* parameter was not sensitive to race manipulations.

Diffusion model in IAT and WIT

To our knowledge, the DM was never applied to data obtained from a Race IAT. However, Klauer Voss, Schmitz and Teige-Mocigemba (2007) have used the DM in two other IAT versions: Flower-Insect and Left-Right political position. In both of these IAT versions, the authors computed a DM with Correct-Incorrect decision thresholds. They interpreted the *Drift-Rates,* as the parameter able to detect the implicit attitude measured by the IAT, i.e., an interference measure that aggregates the information of both accuracy and RT data. Corroborating this view, the authors detected higher *Drift-Rates* in responses to congruent blocks compared to incongruent blocks. The *Threshold Separation* parameter was also interpreted as a measure of response caution. A conservative approach to the task was represented by slowing down response times, to get the correct response. Consistent with this interpretation, the authors found a higher value for *Threshold Separation* on incongruent blocks than on congruent blocks, suggesting that participants were more cautious, when the task was felt as more difficult. The *Nondecision Times* parameter was interpreted as a parameter that captures a set of different processes, such as encoding and the motor execution of the response. This parameter was shown to be higher in the incongruent blocks than in the congruent blocks. The *Relative Starting-Point* parameter was not analyzed in their set of studies.

In another paper, Klauer and Voss (2008), suggested that the DM could also provide information on the processes underlying performance in the WIT. In their view, approaches like the PDP (Jacoby, 1991), that are only based on responses, ignore stereotype effects/bias that can only be observed on RTs. Therefore, the DM, by combining both accuracy and RT data, should be able to detect the initial bias caused by the stereotype activation that happens before the response execution. For example, when primed by a Black portrait, the impression of seeing a Gun, when in the fact a Tool was displayed, would not be captured by accuracy data, because at the moment of response, the impression was already corrected by control processes. This way, the analysis of RTs of incorrect responses on incongruent trials (e.g., Black-Tool) and on congruent trials (e.g., Black-Gun) will be relevant, because the occurrence of an initial bias should lead to faster errors on incongruent trials than on congruent trials.

As a result of these nuances, the authors suggested a DM approach to WIT data slightly different from the one previously used for FPST studies. In fact, they only considered relevant the analysis of the parameters *Drift-Rate* and *Relative Starting-Point.* In their view, the decision thresholds should be set to Gun and Tool (and not correct- incorrect) responses, allowing the *Drift-Rate* parameter

(defined by the differences between Gun-Tools) to function as a measure of discriminability, similar to PDP *C Component*, but incorporating both accuracy and RTs data. A higher *Drift-Rate* on congruent trials (Gun-Black and Tool-White) compared to incongruent trials (Gun-White and Tool-Black), can be interpreted as evidence that the race primes affect the discriminability of the objects (i.e., target). On the other hand, the parameter *Relative Starting-Point* was interpreted as the amount of evidence that participants must accumulate before deciding Gun relatively to Tool decision, i.e., a prior bias of decision (similar to PDP *A Component*). If race leads to a prior decision bias, it is expected that the *Relative Starting-Point* will be higher for trials primed by Black primes, than for trials primed by White primes.

By considering only these two parameters of the DM, authors disregard information that could be provided by the *Nondecision Time* and the *Threshold Separation* parameters. Moreover, this decision made by the authors may have been due to the not contemplation of the possibility that race changes the trade-off between time and decision. Evidence of this effect, as previously mentioned, can be found in Pleskac et al. (2017), in which *Threshold Separation* was higher for Black trials than for White trials, suggesting that time was consumed to give less stereotypically biased responses.

How to use the Diffusion model to model WIT experimental data

What is the best diffusion model to measure performance in a WIT task? Based on both studies with the FPST (Correl et al., 2015; Pleskac et al., 2017) and the article by Klauer and Voss (2008), in WIT tasks, the decision thresholds should be set to Gun and Tool responses. This suggestion differs from the typical Correct and Incorrect decision thresholds. It is important to note that the decision thresholds setting has direct implications on how the parameters should be interpreted (see Table 4). If the decision threshold is set as Gun and Tool, the *Relative Starting-Point* (*zr*) measures a tendency to respond Gun or Tool, while the *Threshold Separation* (*a*) is interpreted as caution to decide between Gun and Tool. However, if we define the decision threshold as Correct and Incorrect decisions, the *Relative Starting-Point* (*zr*) will measure a tendency to respond correctly and the *Threshold Separation* (*a*) will document how cautious participants were between Correct and Incorrect responses. Additionally, while a model with Gun/Tool decisions threshold would permit an analysis of the *Threshold Separation* between Black and White trials as a measure of effort to not use the stereotype (Pleskac et al., 2017), setting the decision threshold as Correct/Incorrect would permit an analysis of the *Threshold Separation* between Incongruent (Black-Tool and White-Gun) and Congruent trials (Black-Gun and White-Tool),

to explore mechanisms in which incongruent information needs more time to be correctly responded (Amodio et al., 2004).

Table 4

Diffusion Model parameters interpretation in each Threshold Decision model

	Threshold Decisions	
	Gun vs Tool (allow to compare Black vs White primes)	**Correct vs Incorrect** (allow to compare Incongruent vs Congruent trials)
***Drift-Rates* (*v*)**	-Measure of response easiness/efficiency in responding to Gun/Tool	- Measure of response efficiency (how easy a correct response is obtained)
***Relative Starting-Point* (*zr*)**	-Tendency to respond Gun/Tool	-Tendency to respond correctly
	-Black vs White comparisons reflects stereotype activation (similar to PDP *A component*)	
***Thershold Separation* (*a*)**	-Cautious to decide between Gun and Tool (more time needed to discriminate between the two responses)	-Cautious to decide between a Correct and an Incorrect response
	-Black vs White comparisons reflects in which Race participants were more cautious	-Congruent vs Incongruent comparisons reflect in which type of information participants were more cautious
***Nondecision Times* (t0)**	- The parameter is unrelated to decision processes	- The parameter is unrelated to decision processes

In conclusion, both approaches to the DM seem to be highly complementary. If Gun-Tool are used as decision thresholds in the DM, *Relative Starting-Point* (*zr*) will provide information on a bias (*zr*) to respond Gun/Tool, which is a relevant measure of stereotype activation (similar to PDP *A Component*). In this model, the *Threshold Separation* (*a*) parameter would capture the participants' caution to respond in Black and White trials. On the other hand, using Correct-Incorrect as decision thresholds, the *Threshold Separation* (*a*) parameter allows access to the participants' caution to correctly respond to congruent and incongruent information, representing a specific effort to cope with

counterstereotypical information at cost of time. As a trade-off for using the Correct-Incorrect decision thresholds, *Relative Starting-Point* (*zr*) becomes just an indicator of the tendency to get a correct response, which is less interesting for the understanding the stereotype bias.

Experiment 1 in light of the Diffusion Model

In this section, our primary goal was to reanalyze the data of experiment 1 (WIT) using the DM, to check if the simple effects detected exclusively in the NPO condition can be better isolated and understood with this approach.

Experiment 1: DM with Correct-Incorrect responses as decision thresholds

In this model, the set of parameters computed were: 2 *Drift-Rates* (Congruent vs Incongruent), 2 *Relative Starting-Points* (Congruent vs Incongruent), 2 *Threshold Separations* (Congruent vs Incongruent), and 2 *Nondecision Times* (Congruent vs Incongruent) parameters. In the Stereotypical congruent trials, Tools were primed by White faces, and Guns primed by Black faces. On the other hand, in the Stereotypical Incongruent trials, Tools were primed by Black faces, and Guns primed by White faces.

These parameters were estimated and tested for their level of fit to the observed data using *fast-dm-30* software (Voss & Voss, 2007). A general index of fit (*p-value*) was provided by the Kolmogorov-Smirnov method and the results subsequently adjusted by the number of task conditions (p-value$^{1/2}$; see Voss & Voss, 2007; Correl, et al., 2015). Higher levels of adjusted fit indexes (probabilities) indicated a good fit of the model. The adjusted fit index was .70(*SD*=.24). Reaching a good adjustment of the model with our data, we tested each DM parameter with a separate Mixed ANOVA model. In this mixed ANOVA, type of trial (Congruent vs Incongruent) was inserted as a within-subjects factor, and Social Condition (NPO and PO) as between-subjects factor.

Drift-Rates(v)-Diffusion Model with Correct-Incorrect thresholds

In this model, *Drift-Rates* (see Figure 12) were interpreted as a measure of response efficiency. We expected to find evidence of stereotype interference when comparing this parameter between Congruent and Incongruent trials. Our prediction was that it should be easier to reach a correct decision in Congruent trials, than in Incongruent trials. However, although our analysis

revealed higher *Drift-Rates* in congruent trials (*M*=2.97; *SD*=1.26), compared to incongruent trials (M=2.82; SD= 1.36), this effect was only marginally significant, F (1,123) =1.26, p=.09, η^2=.02.

The Block effect was not moderated by Social condition, F <1 and the Social Conditions main effect was also not significant, F (1,123) =2.00, p=.16, η^2=.02. (see Figure 12), indicating that there was no SF effect in this parameter.

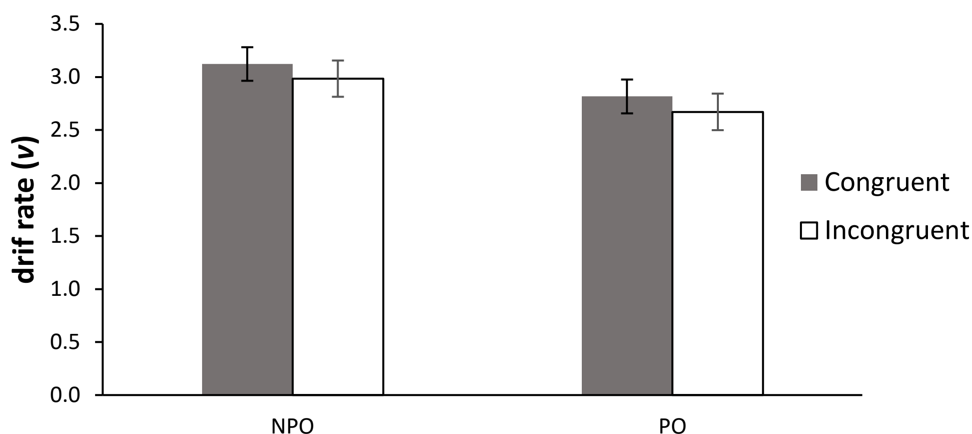

Figure 12. *Drift-Rates for WIT trials (Congruent vs Incongruent) in both Social Conditions (NPO and PO). Errors bars denote one standard error around the mean.*

Relative Starting Point(zr)-Diffusion Model with Correct-Incorrect thresholds

In this model, the *Relative Starting-Point* (see Figure 13) represented the participants' tendency to get the Correct/Incorrect response. Values higher than .5 indicate a tendency to get the correct response, indicating that less information/evidence is needed to get a correct response. In contrast, values lower than .5 indicate a tendency to get an incorrect response, indicating that less information /evidence is needed to get an incorrect response. In general, our participants presented a relative starting-point of .46 (*SD*=.13), which indicates that more evidence was required to reach a correct response than to reach an incorrect response.

A stereotype effect in this parameter would be reflected by a higher *Relative Starting-Point* for Congruent trials compared to Incongruent trials. However, the main effect of Type of Trial was not significant, F (1,123) =1.14, p=.29, η^2=.01. Moreover, no significant main effect for Social Condition, F<1, or Social Condition x Type of Trial interaction was found, F<1.

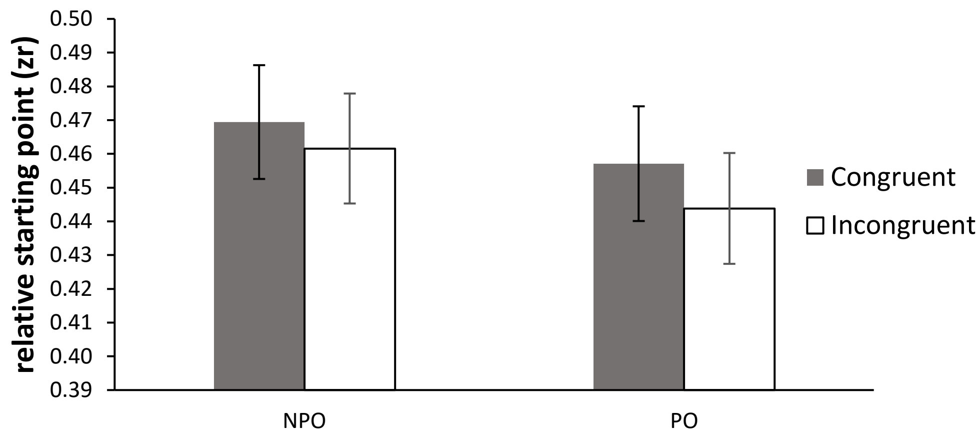

Figure 14. *Relative Starting-Points for WIT trials (Congruent vs Incongruent) in both Social Conditions (NPO and PO). Errors bars denote one standard error around the mean.*

Threshold Separation (a)-Diffusion Model with Correct-Incorrect thresholds

Threshold Separation (see Figure 14), represents the trade-off between time and response, i.e., higher values indicate that participants were more cautious (spent more time) in order to get the correct response. A stereotype effect in this parameter should be observed in the differences between Incongruent and Congruent trials, so that a higher threshold separation in Incongruent trials than in Congruent trials, would indicate an effort to avoid be stereotype bias.

Although we did not detect the main effect of Type of Trial (F<1), we have found a marginal interaction between Type of Trial and Social Condition, F (1,123) =3.50, p=.06, η^2=.03. Simple effects analysis shows that in PO there was no difference between Congruent (M=.66; SD=.12) and Incongruent trials (M=.64; SD=.11), (t<1), suggesting that participants in this condition used the same level of caution independently of the Type of Trial. For participants in NPO, there was a marginal difference between Congruent (M=.61; SD=.12) and Incongruent (M=.63; SD=.11) trials, t(123)= 3.06; p=.08, d=.18. This result indicates that participants in NPO, spent more time in Incongruent trials in order to achieve the correct response threshold than in Congruent trials.

We have also detected a marginal effect of Social Condition, *F* (1,123) =3.00, *p*=.09, η^2=.02, indicating a tendency for a higher *Threshold Separation* in PO (*M*=.65; *SD*=.10) than in NPO (*M*=.62; *SD*=.10), suggesting that participants in PO were more cautious in order to achieve the correct response threshold, than those in NPO.

In chapter 3, our first analyses showed that participants in the PO were faster and more prone to errors (compatible with an impulsive approach) than those in the NPO condition. Yet, the *Threshold Separation* analysis suggests those in PO were more cautious, requiring more time to reach the correct response threshold. This apparent incongruent result could be explained by participants in PO requiring more time to achieve the correct response, to compensate their increased proneness to error compared to NPO.

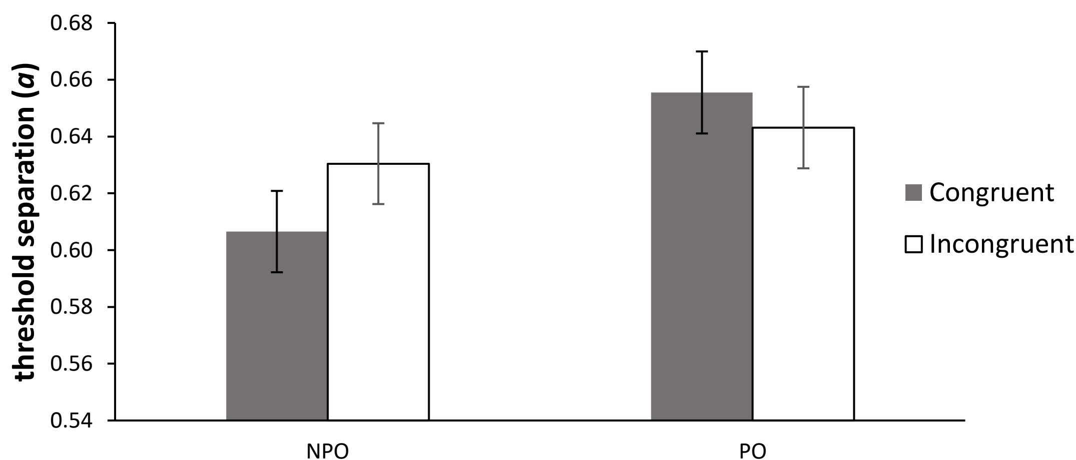

Figure 15. *Threshold Separation for WIT trials (Congruent vs Incongruent) in both Social Conditions (NPO and PO). Errors bars denote one standard error around the mean.*

Nondecision Time (t0)-Diffusion Model with Correct-Incorrect thresholds

Nondecision Times (see Figure 15), represent the time spent in every process that was not related to the decision processes (e.g., response execution, information encoding, switching cost…). Because this parameter is unrelated to decision processes, we did not expect effects related to Type of Trial (i.e., stereotype effects). Our results were congruent with this hypothesis, since there was no significant main effect of Type of Trial (*F*<1) or a Type of Trial x Social condition interaction (*F*<1).

We have found a main effect of Social Condition, F (1,123) =6.20, p=.02, η^2=.05, showing that participants in PO (M=306.97; SD=67.84) had faster *Nondecision Times* than participants in NPO (M=337.50; SD=67.86).

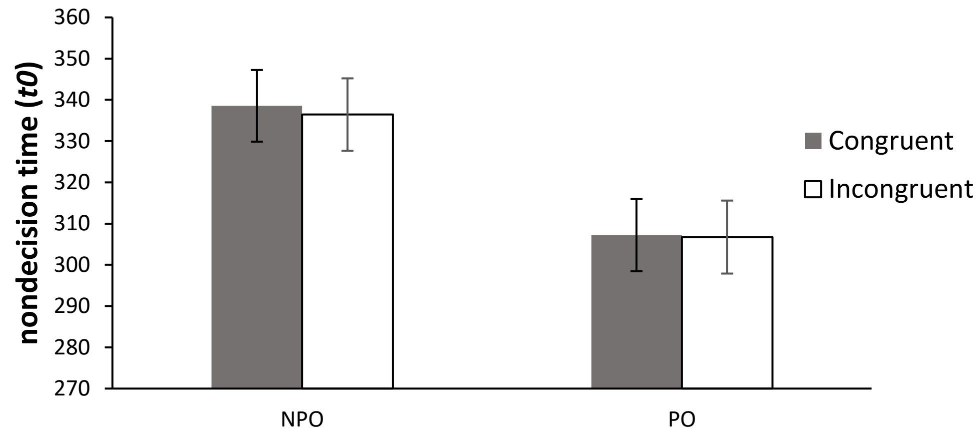

Figure 16. *Nondecision Times for WIT trials (Congruent vs Incongruent) in both Social Conditions (NPO and PO). Errors bars denote one standard error around the mean.*

Summary of DM with Correct-Incorrect responses as decision thresholds

DM parameters associated with correct-incorrect responses of experiment 1 data revealed that: a) stereotype information (incongruent vs congruent) was likely to be affecting participants' efficiency to WIT trials (lower drift-rates for incongruent trials than congruent trials) in both PO and NPO conditions; b) whereas participants in PO needed to be more cautious to provide a correct response, those in NPO conditions were more cautious in incongruent trials, than in congruent trials (*Threshold Separation* was marginally higher for incongruent trials, than for congruent trials only in NPO); c) *Nondecision Times* were faster for those in PO; and d) there was no effect related to *Relative Starting-Point*.

Experiment 1: DM with Gun-Tool responses as decision thresholds

For this model, we computed 4 *Drift-Rates* (White-Tool, Black-Tool, White-Gun and Black-Tool trials), 2 *Relative Starting-Points* (Black and White trials), 2 *Threshold Separations* (Black and White Trials), and 4 *Nondecision Times* (White-Tool, Black-Tool, White-Gun and Black-Tool trials) parameters.

These parameters were estimated and tested for their levels of fit to the observed data using *fast-dm-30* software (Voss & Voss, 2007). A general index of fit (*p-value*) was provided by the Kolmogorov-Smirnov method. The results were subsequently adjusted by the number of task conditions (p-value$^{1/4}$; see Voss & Voss, 2007; Correl et al., 2015). Higher levels of adjusted fit indexes (probabilities) indicated a good fit of the model. The adjusted fit index was .74(*SD*=.16). Having reached a good adjustment to our data, we analyzed each DM parameter using separate Mixed ANOVAs.

Drift-Rates(v)-Diffusion Model with Gun-Tool thresholds

Since Tool was set as the bottom decision threshold, the *Drift-Rates* values for these decisions were negative. Therefore, to directly compare the *Drift-Rates* for Tool and Gun, values for Tool were inverted. The data was then tested using a Mixed ANOVA with Target (Tool vs Gun) and Prime (Black and White) as a within-subjects factor, and Social Condition (NPO vs PO) as a between-subjects factor (see Figure 16).

The analysis revealed a main effect of Prime, F (1,123) =10.89, p<.01, η^2=.08, indicating that it was easier to respond in trials primed by White portraits (M=2.54; SD=1.36), than in trials primed by Black portraits (*M*=2.41; *SD*=1.31). A significant main effect of Target, F (1,123) =18.45, p<.001, η^2=.13, indicates that it was easier to make a decision for Gun (*M*=2.65; *SD*=1.39), than for Tool (*M*=2.33; *SD*=1.36). We have also found an interaction for Prime x Target F (1,123) =13.51, p<.02, η^2=.10. This interaction suggests that it was more difficult to make a decision for Tool when the target was primed by Black portraits (*M*=2.15; *SD*=1.44), than when primed by White portraits (*M*=2.50; *SD*=1.41), t(123) =18.28, p=<.001, d=.25 . When the target was Gun, the easiness of the decision was not significantly different between White (*M*=2.63; *SD*=1.50) and Black portraits (*M*=2.67; *SD*=1.38), t<1. Together, these results suggest that a stereotype effect in terms of response easiness (time and accuracy) was promoted by the difficulty of responding correctly to Black-Tool.

Although we detected a Stereotype effect on *Drift-Rates*, we did not find evidence of a moderation by Social Condition (*F*<1). Moreover, no other effect was significant, including the Social Condition main effect, F (1,123) =1.84, p=.18, η^2=.02; Target x Social Condition interaction, F (1,123) =1.34, p=.25, η^2=.01; and the Prime x Social Condition interaction, *F*<1.

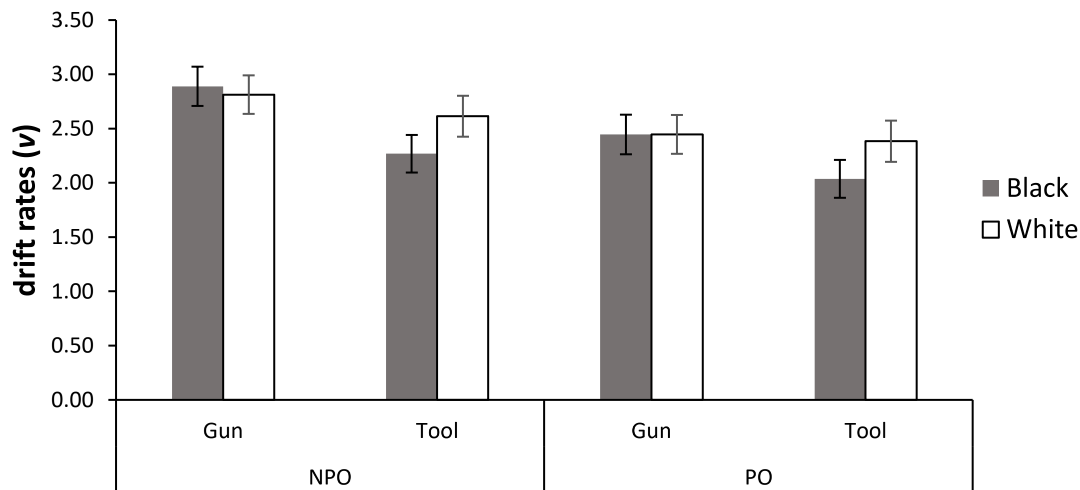

Figure 17. *Drift-Rates for WIT trials (for each type of trial) in both Social Conditions (NPO and PO). Errors bars denote one standard error around the mean.*

Relative Starting-Point(zr)-Diffusion Model with Gun-Tool thresholds

Relative Starting-Point represents the participants' prior tendency to respond Gun/Tool. In this parameter, values higher than .50 indicate a tendency to respond Gun, while values lower than .50 represent a tendency to respond Tool. Our participants had in average a *Relative Starting-Point* of .47, which can be interpreted as a general tendency to respond Tool. Since the response key for Tool was on the right side of the keyboard, this result could have been caused by the participants' tendency of using their dominant hand (usually the right one) to respond.

The mixed ANOVA with Prime (Black vs White) as within-subjects factor, and Social Condition (NPO vs PO) as between-subjects factor (see Figure 17), did not detect the main effect of Prime (*F*<1), suggesting that there was no stereotype effect in this parameter. The main effect of Social Condition (*F*<1) and Prime x Social Condition interaction (*F*<1) were also not significant.

These results suggest that stereotype information was not affecting a possible prior bias of the participants response in a specific direction (Tool or Gun).

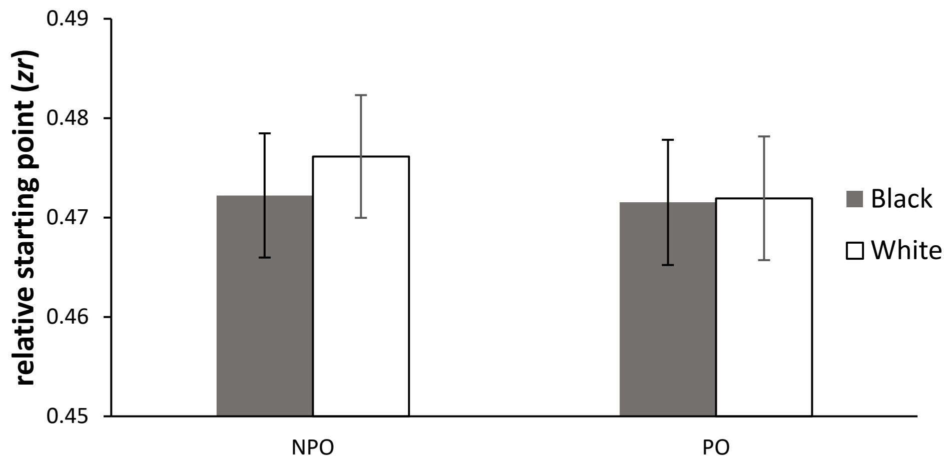

Figure 18. *Relative Starting-Point for WIT trials (by Black and White primes) in both Social Conditions (NPO and PO). Errors bars denote one standard error around the mean.*

Threshold Separation(a)-Diffusion Model with Gun-Tool thresholds

Higher values of *Threshold Separation* suggest that more information was needed to discriminate between Gun and Tool decisions. In this parameter, lower values indicate that individuals needed less information for their responses. A Mixed ANOVA with Prime (Black vs White) as a within-subjects factor, and Social Condition (NPO vs PO) as a between-subjects factor (see Figure 18) with *Threshold Separation* values as a dependent variable was conducted. The main effect of Prime was not significant, $F<1$, suggesting that participants were not more cautious to respond after being primed with either race.

The main effect of Social Condition, $F\,(1,123) = 5.87$, $p=.02$, $\eta^2=.05$, indicates that participants in PO (M=.68; SD=.07) had a higher *Threshold Separation* than participants in NPO (M=.65; SD=.07). This effect reinforces our previous findings (the DM analysis with Correct-Incorrect thresholds) that participants in PO were more cautious in their responses, than those in NPO (i.e., they needed more time to respond).

The interaction Prime x Social Condition was not significant, F (1,123) =2.13, p=.15, η^2=.02, suggesting that the differences in *Threshold Separations* in our Social Conditions were not specifically related to the Primes.

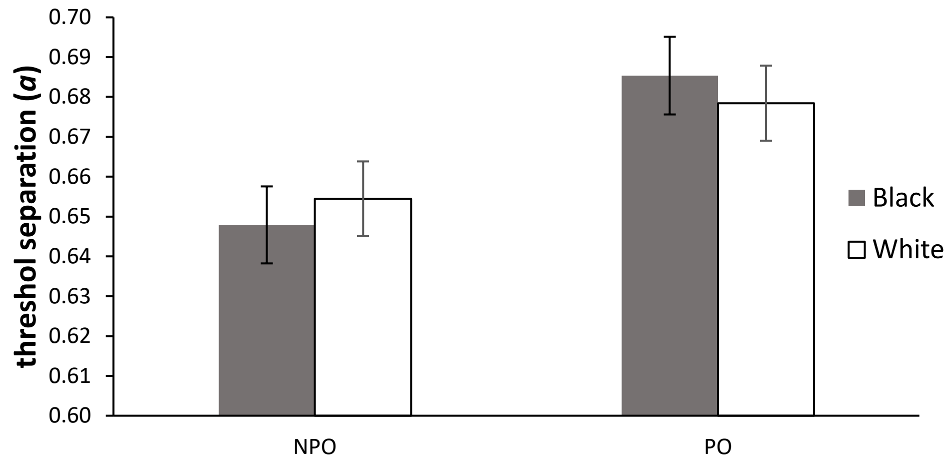

Figure 19. *Threshold Separation for WIT trials (by Black and White primes) in both Social Conditions (NPO and PO). Errors bars denote one standard error around the mean.*

Nondecision Times(t0)-Diffusion Model with Gun-Tool thresholds

Nondecision Times (see Figure 19) represents the time spent in every process that was not related to the decision processes (e.g., response execution, information encoding, switching cost…). In the mixed ANOVA for *Nondecision Times*, we inserted Target (Tool vs Gun) and Prime (Black vs White) as within-subjects' factors and the Social Condition (NPO vs PO) as a between-subjects factor. This analysis revealed no effect of stereotype on *Nondecision Times*, Target x Prime interaction F<1, which is compatible with an interpretation of stereotype effects as effects on decisional processes and not of response competition.

We have found a main effect of Target, F (1,123) =4.34, p=.04, η^2=.03, indicating that participants were faster at Gun trials (M=315; SD=67.69) than at Tool trials (M=319.62; SD=68.08). The main effect of Social Condition was also significant, F (1,123) =6.47, p=.01, η^2=.05, suggesting that the participants in PO (M=302.48; SD=66.52) were generally faster in their responses, than those in NPO (M=333.06; SD=66.54).

The Prime main effect ($F<1$); Target x Social Condition interaction ($F<1$); Prime x Social Condition interaction, F (1,123) =1.19, p=.28, η^2=.01; and the Target x Prime x Social Condition interaction ($F<1$) were all not statistically significant.

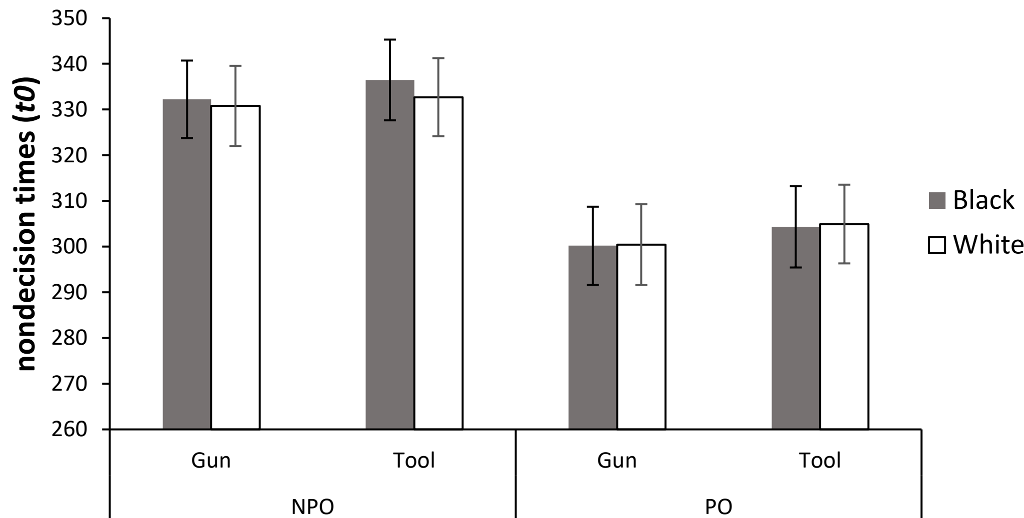

Figure 110. *Nondecision times for WIT trials (by Black and White primes) in both Social Conditions (NPO and PO). Errors bars denote one standard error around the mean.*

Summary of DM with Gun-Tool responses as decision thresholds

The DM parameters associated with the Gun-Tool responses of experiment 1 data revealed: a) a stereotype effect, associated with participants finding more difficult to respond to the Black-Tool trials, than other trials (as shown by the *Drift-Rates*), that was not moderated by Social Conditions; b) neither Social Condition impacted the participants general tendency to respond Tool (suggested by the analysis of the *Relative Starting-Point*); c) participants were not more cautious to respond after a Black or a White prime (contrary to Plesack et al. (2017) with the FPST), but participants in PO required more time to discriminate between Gun and Tool responses, when compared to NPO; d) *Nondecision Times* were faster for Gun responses, which may have been influenced by the keyboard position for the Gun response. Participants in PO responded faster than those in NPO, replicating the effect detected with RTs.

DM in Experiment 1 – Final Discussion

To disambiguate different components of the decisional process (e.g., efficiency, decision bias, caution and variables related to response execution), we have applied the DM to the data of our study 1. To the best of our knowledge, no previous research has used the DM to analyze WIT data. To meet this challenge, we explored two different approaches for the implementation of DM in WIT data. A more classical/standard approach that used Correct and Incorrect as decision thresholds (similar to Klauer et al., 2007) and an alternative, based on the FPSTs (Correl, et al., 2015; Pleskac, et al., 2017), that sets Gun and Tool as decision thresholds (reflecting the Shoot/No shoot used in FPST). As summarized in Table 4 the parameters' meaning and application differs depending on the decision's threshold used in the DM.

Using Correct-Incorrect as the decision threshold in the DM, we have found evidence of a stereotype effect for the parameter *Drift-Rate*. In this parameter, participants were more efficient in Congruent trials that in Incongruent trials, suggesting that counterstereotypical information reduces the participants' capacity to discriminate the targets. However, we have not found an effect for social condition for this parameter.

Participants in PO seemed to need to be more cautious (*Threshold Separation*) when reaching a correct response, than those in NPO. On the other hand, participants in NPO seemed to search for more evidence in Incongruent trials than in Congruent trials. This pattern matches the results we have previously reported in Chapter 3 for the analysis of RT of correct responses. Together, the two analyses suggest that Social Conditions modulated the way the participants coped with stereotype information.

The DM with Gun-Tool as the decision thresholds enabled us to find clear evidence of a general stereotyping effect. The analysis of *Drift-Rates* showed that it was harder for participants to respond to the Black-Tool trials. We can then hypothesize that the stereotype association Black-Gun was activated in the participants' minds, making it more difficult to give correct response in Black-Tool trials. Social Conditions also lead to differences in the participants decisional process, with those in PO searching for more evidence in order to discriminate between the two responses (*Threshold Separation*).

The analyses of both models (Correct-Incorrect Threshold and Gun-Tool Threshold) suggest that stereotype activation was similar for the two social conditions, but participants tended to exert different types of control over their responses. Whereas in PO, participants were more cautious for

all responses (even if more prone to error), those in NPO showed to be more cautious only when dealing with counterstereotypical information. Moreover, the threshold difference between Incongruent and Congruent trials was associated with less stereotype bias in ERs ($r(125)=.20$, $p=.02$), suggesting that participants in NPO were using a type of control that aimed to discriminate between stereotypical and counterstereotypical information. This control mechanism seems to be similar to what was described by Amodio et al. (2004), in which the slower responses were given on trials with counterstereotypical information.

We have not found an effect for the *Relative Staring-Points*, suggesting that the participants' prior bias was not influenced by stereotypical information or by our manipulation of Social Conditions. Overall, these values simply indicate a slight tendency of participants to respond Tool, possibly because it captures the participants' tendency of using their dominant hand (usually the right one).

The *Nondecision Times* analysis showed that, in PO conditions, participants were faster in their responses, replicating Social Facilitation effects (e.g., Triplett, 1898). Importantly, these faster responses were not associated with the decisional task that they were performing. This effect could have simply resulted from the participants' being more driven to perform the task, which is consistent with what has been advocated since the beginning of the study of SF (e.g., Triplett, 1898; Dashiell, 1930; Cottrell, 1968).

In sum, the application of the DM to our data permitted the disambiguation of the decisional processes, by simultaneously incorporating the data of the two response components (time and accuracy). The analysis performed with DM clarified that stereotype activation was not moderated by the presence of others, even though the participants in the presence of others were faster in their responses, committed more errors and required more information to achieve a correct response.

Experiment 3- Implicit Association Test with restriction of time

The use of DM in Experiment 1 was critical for the detection and understanding of possible hidden SF effects. In experiment 2, we have found evidence of a stronger IAT effect (bias) on the slower responses of participants in the PO condition. However, the IAT data of experiment 2 lacks sufficient response error variability to meet the necessary requirements of a DM analysis. To deal with

this issue, in experiment 3, we increased the number of trials of the IAT task and added a response time restriction.

As Klauer, et al. (2007), we computed the DM parameters in function of the decision thresholds Correct-Incorrect. The parameters *Drift-Rates, Relative Starting-Point, Threshold Separation* and *Nondecision Times* were calculated for each block (Incongruent vs Congruent); see Table 5.

The interpretation of the parameters also follows Klauer et al. (2007). *Drift-Rate* will be used as a measure of performance (efficiency) that combines RTs and Accuracy data, and *Threshold Separation* will be interpreted as a measure of the participants' caution while performing the task. Finally, *Nondecision Times* reflects the general time spent completing the task. Although, Klauer, et al. (2007) considers that the *Relative Starting-Point* does not have a meaningful interpretation in their versions of the IAT, we believe that, in the Race IAT, this parameter can be interpreted as a measure of prior Stereotype expectations. If our interpretation is correct, we expect a higher *Relative Starting-Point* on the Congruent block, than on the Incongruent block.

Table 5

Diffusion Model parameters interpretation for IAT

	Threshold Decisions
	Correct vs Incorrect
Drift-Rates (*v*)	- Measure of response efficiency (how easy a correct response is obtained)
Relative Starting-Point (*zr*)	- Prior expectation lead to higher values in congruent trials
Threshold Separation (*a*)	-Level of caution to decide between Correct and Incorrect responses (Congruent vs Incongruent trials reflect the relative effort to suppress stereotype bias at the cost of time)
Nondecision Time (t0**)**	-The parameter is unrelated to decision processes

Experiment 3- Methods

Participants and Design

A total of 100 non-black female subjects with an average age of 20.27 (*SD*=4.72) enrolled at ISPA, volunteered and participated in this experiment to obtain course credits. As in study 2, Social Condition (NPO vs. PO) was manipulated between participants and type of Block (Incongruent Vs Congruent) was manipulated within-subjects.

Procedure

After the subjects gave their informed consent to participate in an evaluation of words and images experiment, experimental sessions were scheduled in group (PO) or individually (NPO) according to the randomly assigned social condition. In the NPO condition, each participant completed the IAT in an experimental room, without the presence of the experimenter. In the PO condition, participants arrived in groups of 4 to 8 individuals and performed the IAT at the same time, and in the same experimental room.

As in experiment 1 and 2, the experimenter was present in the PO condition, but positioned in a way that it was not possible to monitor the participants' computer screens. Task instructions were provided on the computer screen and informed the participants that they would have to evaluate a set of words as Positive or Negative and a set of faces as Black or White.

In this experiment, we have used a modified version of the IAT of experiment 2 in order meet the DM data variability requirements on Accuracy (correct and incorrect responses) and RT. For this purpose, we have added a response time restriction of 600 ms, similar to the 550 ms used in WIT (Payne, 2001) and increased the number of trials to 180 in each of the critical blocks (Congruent and Incongruent). Moreover, we have changed the order of the critical blocks (Block 3- Congruent and Block 5- Incongruent), to rule out the possibility of order effects in our data. The number of trials on non-critical blocks (1, 2 and 4) was set to 60 trials per block.

In the first block, participants evaluated 60 faces (30 black and 30 white faces) as black or white. In the second block, participants evaluated 60 word as bad (infection, catastrophe, expulsion, ambush, degraded, corpse, wound, coffin, poison, bomb, rubble, avalanche, snot and maggot) and good (hug, friend, love, kiss, child, embryo, fairy, fertile, flower, harp, oasis, paradise, tenderness, truth, virtue). The third block was the first critical block (Congruent) and participants had to evaluate the

same faces and words, but with the categories Black and Negative on the same side of the screen and the categories White and Positive on the other side. Fourth block was equal to the first block, but the position of the category Black and White was changed, as in experiment 2. The last block, the Incongruent Block (critical), was similar to the third block, but the Black and Positive categories were grouped on the same side, and White and Negative were placed on the opposite side of the screen.

Dependent Variables

To analyze experiment 3, we extracted two sets of dependent variables from the critical blocks (Block 3- Congruent; Block 5- Incongruent), reflecting the traditional measures used in IAT research and our current approach of using CDF-plots and DM parameters.

The **d600 index** is a standard index used in IAT studies. This index was computed following Greenwald, Nosek and Banaji (2003): 1-Eliminate trials with RT greater than 10,000 ms and subjects with more than 10% of trials with RT less than 300 ms; 2-Compute the mean of correct RT for each critical block (i.e., Congruent and Incongruent) and the standard deviation of the two blocks together; 3-Replace each error RT with the block mean and add 600 ms; 4-Compute the average of the critical blocks and the difference between them; 5-Divide the difference by the standard deviation of the two blocks. The completion of these steps ensures a *d600* index for each participant. A higher *d600* value indicates a more implicit negative attitude towards black people.

Error Rates (ERs) were calculated for each Block (Incongruent and Congruent). Participants' errors were divided by the total number of responded trials. Trials with no responses were replaced by missing values and were not used for ERs computation.

PDP Components were computed based on Stewart, von Hippel and Radvansky (2009). The PDP *C Component* (probability of correctly evaluating the target) was calculated by subtracting the probability of correct responses on Congruent trials by the probability of an incorrect response on Incongruent trials. The PDP *A Component* (probability of responding in a stereotypical way when control fails) was calculated by the probability of an incorrect response in Incongruent trials divided by (1-C).

Reaction Times of Correct Responses (RTs) was computed as an average of correct RTs for each block, which resulted in 2 RTs averages (congruent and incongruent) for each participant.

CDF plots were computed for each participant, the RT distributions of correct responses in the Incongruent and Congruent blocks were separately split in 6 bins. For each corresponding bin, we computed the difference between the Incongruent and Congruent RT (i.e., $Bin1_{Incongruent}$-$Bin1_{Congruent}$; … ;$Bin6_{Incongruent}$-$Bin6_{Congruent}$). Higher values in the CDF indicates a stronger stereotype bias.

DM parameters with Correct-Incorrect responses as decision thresholds were applied as in Klauer, et al. (2007), permitting access to *Drift-Rates, Relative Starting-Point, Threshold Separation and Nondecision Times* for each block (Incongruent vs Congruent). As in experiment 1, the DM computation was performed using the *fast-dm-30* software (Voss & Voss, 2007). The general adjusted index of fit was *M*=.70 (*SD*=.22), which can be interpreted as a good fit.

Experiment 3- Results

Preliminary analysis revealed that two participants had an abnormal number of missing values in critical blocks (307 and 311 in a total of 360 trials) and that three participants had 10% of their RTs under 100 ms (exclusion criteria for IAT by Greenwald, Nosek & Banaji, 2003). These five participants were excluded from the sample and the analysis was performed with 95 participants (44 in NPO and 51 in PO).

d600

As expected by adding a response time restriction and trials to the IAT, participants responses were shown to be more sensitive to the Stereotype Bias increasing the *d600* values in this experiment (*M*= .55; *SD*=.49) compared to what we reported in experiment 2 (*M*=.20; *SD*=.33). We have also found that participants in the PO (M_{PO}= .63; *SD*=.54) stereotyped more than those in the NPO condition (M_{nPO}= .45; *SD*=.39; t(93) =-1.84, *p*=.03, *d*=.37), see Figure 20. This result replicates our findings (experiment 2) and oppose the findings of Castelli and Tomelleri (2008), that reported lower stereotype for participants in the presence of others.

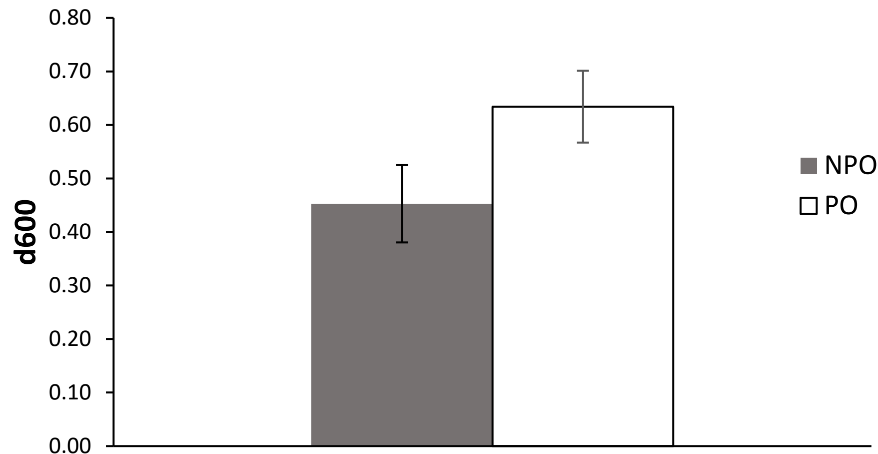

Figure 20. *d600 index in NPO and PO. Errors bars denote one standard error around the mean.*

Error Rates

Using a mixed ANOVA with Block (Incongruent vs Congruent) as a within-subjects factor and Social Condition (NPO vs PO) as a between-subjects factor (Figure 21) we detected a main effect for Block, F (1,93) =93.83, p<.001, η^2=.50, indicating that the participants ERs in the Incongruent block (M=.16;SD=.09) was higher, than in the Congruent block (M=.10;SD=.05).

However, the main effect of Social Condition, F (1,93) =2.06, p=.16, η^2=.02, and the Social Condition x Block interaction, F (1,93) =1.74, p=.19, η^2=.02 were not significant. These results indicate that Social Conditions did not affect general ERs, and did not moderate the stereotype effect (Block main effect).

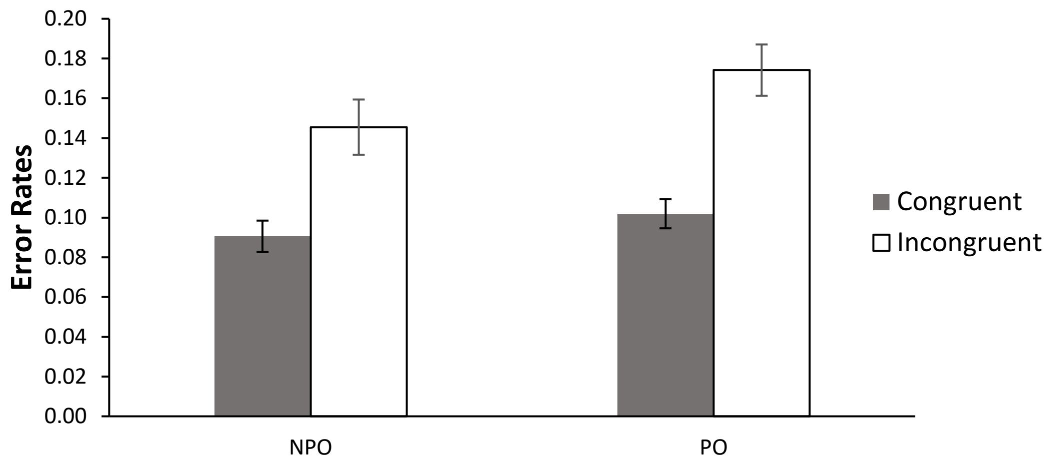

Figure 21. *Error Rates for Blocks (Incongruent vs Congruent) in both NPO and PO. Errors bars denote one standard error around the mean.*

Reaction time of Correct responses

RTs were analyzed by using a mixed ANOVA with Block (Incongruent vs Congruent) as a within-subjects factor and Social Condition (NPO vs PO) as a between-subjects factor, see Figure 22. No significant effects were detected on this analysis; Block effect (F<1), Social Condition effect (F<1), interaction Block x Social Condition, F (1,93) =1.01, p=.32, η^2=.01.

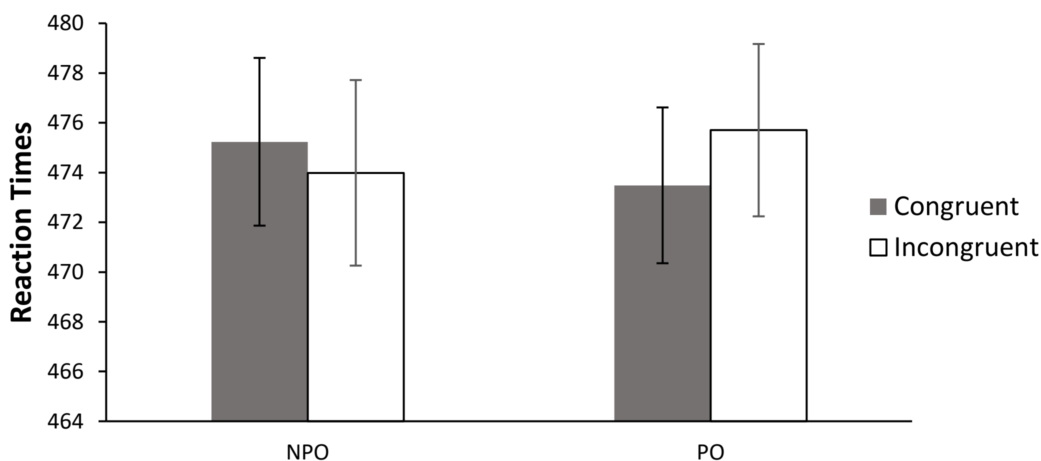

Figure 22. *Reactions times for Blocks (Incongruent vs Congruent) in both NPO and PO. Errors bars denote one standard error around the mean.*

PDP Components

The Control Component was .75, see Table 6, which was lower than what was observed in experiment 2 (.90). This could be explained by an increase of task demands, resulting from the response time restriction implemented in experiment 3 which could have hindered the participants' performance.

Differences between NPO and PO in the Control and Automatic components were tested using two separate t-tests. No differences were detected for the Control component (t<1) or for the Automatic component (t<1).

Table 6

PDP components likelihoods (SDs) regarding Social Conditions

Components	**Social Condition**		**Overall**
	NPO	*PO*	
C	.76(.14)	.74(.13)	.75(.13)
A	.39(.10)	.38(.11)	.38(.10)

Cumulative Distribution Frequency

Two CDF analysis were performed. The first analysis replicates the approach followed in experiment 2 and used RTs as dependent measure. The second CDF analysis used the ERs as a dependent measure.

CDF plots for Incongruent-Congruent differences (see Figure 23) were analyzed through a mixed ANOVA with Social Condition as between-subjects factor and Bins (6) as within-subjects factor. The ANOVA detected a main effect of Bins, F (5,465) =16.83, p<.001, η^2=.15, as in the experiment 2. This effect indicates that the stereotype bias increases over the time. Contrary to our findings in experiment 2, bias seems to be decreasing over the last bins. A closer look at Figure 23, also reveals that stereotype bias values were negative for the first bins, indicating, that participants responses were faster in the Incongruent block. Since the Congruent Block was performed before the Incongruent Block, it is possible that participants were still not sufficiently trained in using the response keys, leading to slower RT within the Congruent block. The negative values on the first bins could then possibly be explained by learning/order effects.

The overall analysis did not detect effects promoted by Social Condition, F (1,93) =1.14, p=.29, η^2=.01, and the interaction Social Condition x Bins was also not significant (F<1). However, Figure 23, shows that in the last bin (i.e., participants' slowest responses), Social Condition moderated the stereotype effect. While the participants in NPO have a stereotype bias lower than 0 (M=-1.13; SD=11.70), those in PO have positive values on the Stereotype index (M=4.63; SD=11.60). This difference was statistically significant, t(93)=-2.52, p=.02, d=.50.

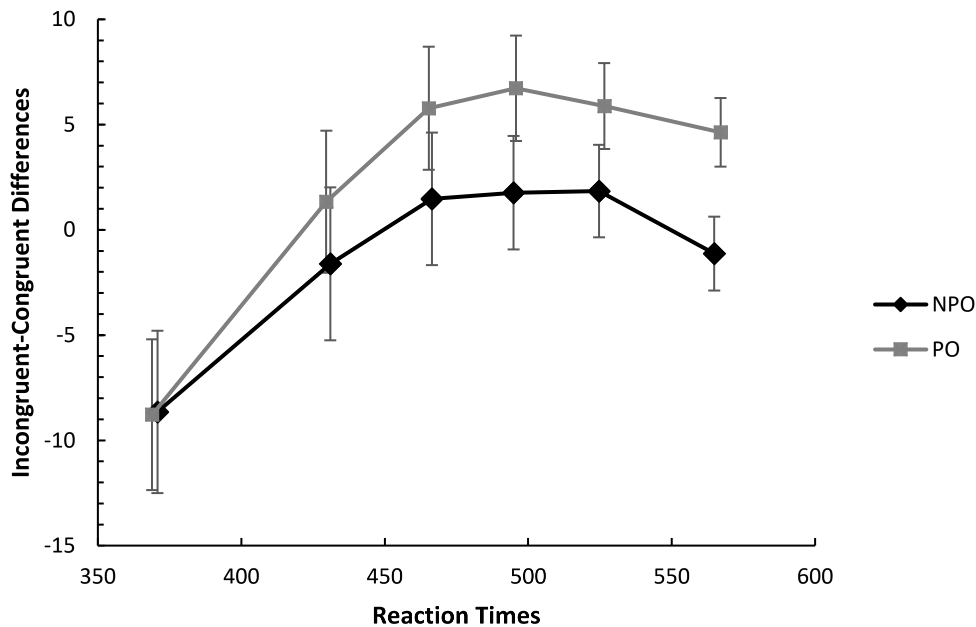

Figure 23. *CDF Plots with Incongruent-Congruent differences for Social Condition (NPO vs PO). Each marker represents a bin which is associate in a Reaction Time Average. Errors bars denote one standard error around the mean.*

Using the ER as a dependent variable, the CDF plots for Incongruent-Congruent differences (see Figure 24) were analyzed with a mixed ANOVA with Social Condition as between-subjects factor and Bins (6) as within-subjects factor. The ANOVA detected a main effect of Bins, F (5,465) =4.59, p<.001, η^2=.05, indicating, as in experiment 2, that stereotype bias (now based on ERs) increased over the time.

Congruent with what was previously reported, we have not found a main effect of Social Condition, F (1,93) =1.77, p=.19, η^2=.02., or a Bins x Social Condition interaction (F<1).

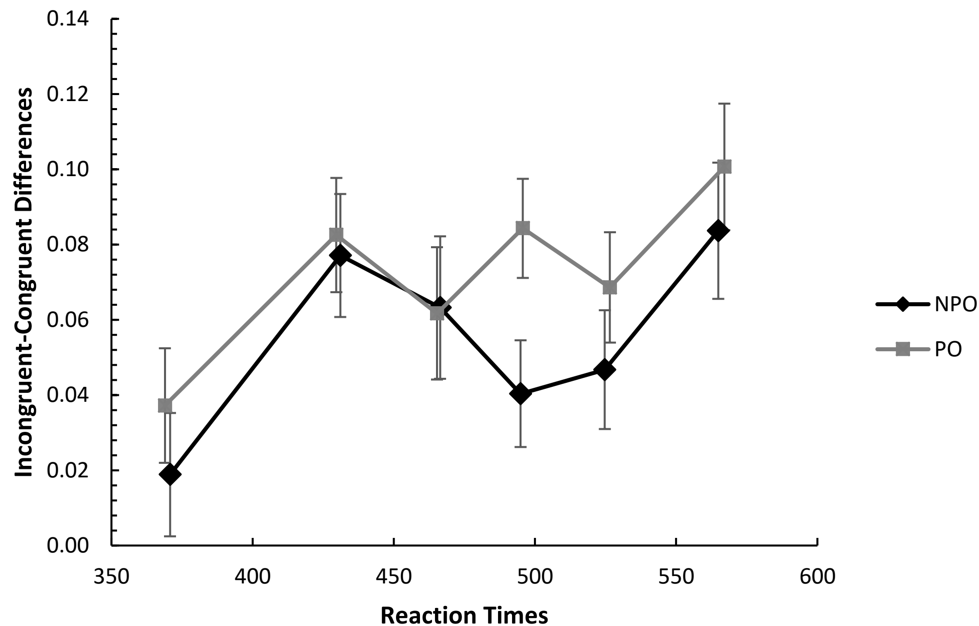

Figure 24. *CDF Plots with Incongruent-Congruent differences for Social Condition (NPO vs PO). Each marker represents a bin which is associate in a Reaction Time Average. Errors bars denote one standard error around the mean.*

Diffusion Model with Correct-Incorrect thresholds Results

Each parameter was analyzed with a mixed ANOVA with Block (congruent versus Incongruent) as within-subjects factor and Social Condition (PO vs NPO) as between-subjects factor.

Drift-Rates(v)

Drift-Rates can be interpreted as a measure of performance that takes into account RT and Accuracy data (i.e., efficiency). Thus, it is expected lower *Drift-Rates* in the Incongruent block compared to the Congruent block, consistent with the typical stereotype effect. The main effect of Block, F (1,93) =26.50, p<.001, η^2=.22, indicates that the Incongruent block (M=3.71; SD=1.19) had lower values than the Congruent block (M=4.38; SD=0.82), confirming that participants were more efficient in dealing with stereotypical information than with counterstereotypical information. (see Figure 25).

The interaction Block x Social Condition was marginal, F (1,93) =3.43, p=.07, η^2=.04. This marginal effect is important to our aims because it suggests that participants in PO are more prone to the stereotype effect, i.e., they have more trouble dealing with the counterstereotypical information.

The main effect of Social Condition was not significant (F<1).

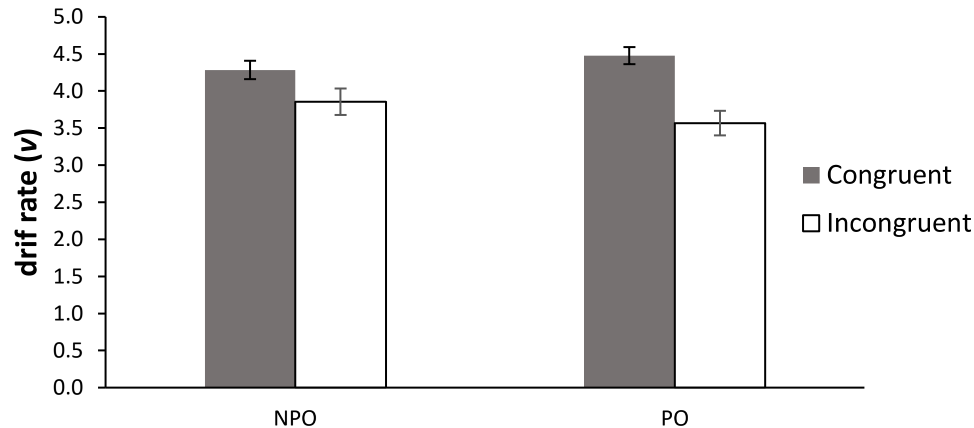

Figure 25. *Drift-rates for IAT trials (Congruent vs Incongruent) in both Social Conditions (NPO and PO). Errors bars denote one standard error around the mean.*

Relative Starting-point(zr)

The mixed ANOVA revealed a main effect of Block, F (1,93) =9.13, p<.01, η^2=.09, indicating that the Congruent block (M=.53; SD=.18) have higher *Relative Starting-Points* than the Incongruent block (M=.48; SD=.15). This suggests that stereotype was biasing participants wrong decisions in the Incongruent block, while in the Congruent block participants were biased to correct decisions.

Although graphically (see Figure 26) the stereotype influence seemed to be higher in NPO, the interaction Block x Social Condition did not reach statistical significance, F (1,93) =2.60, p=.11, η^2=.03.

The main effect of Social Condition was also not significant, F (1,93) =2.33, p=.13, η^2=.03.

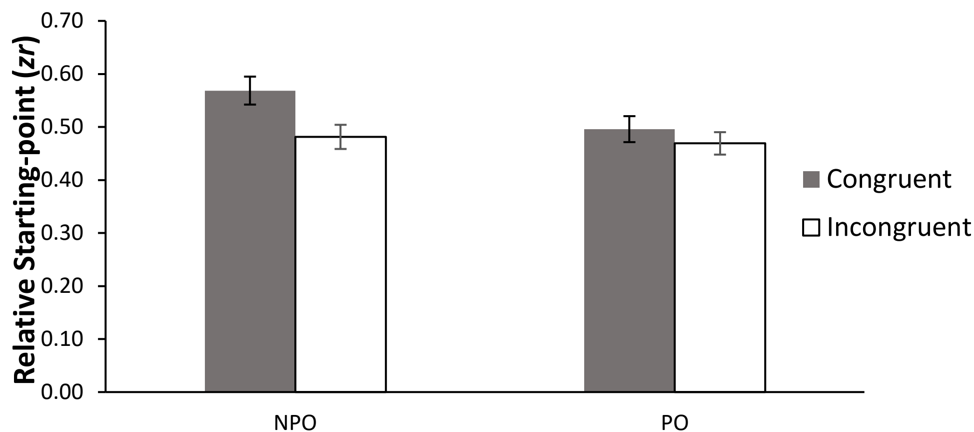

Figure 26. *Relative Starting-points for IAT trials (Congruent vs Incongruent) in both Social Conditions (NPO and PO). Errors bars denote one standard error around the mean.*

Threshold Separation(a)

Analysis revealed a main effect of Block, F (1,93) =7.05, p<.01, η^2=.07, suggesting that participants were more cautious to give correct responses in the Congruent block (M=.69; SD=.18) than in the Incongruent block (M=.65; SD=.12), see Figure 27.

No other effects were significant, including the Social condition main effect (F<1) and the Block x Social Condition interaction (F<1).

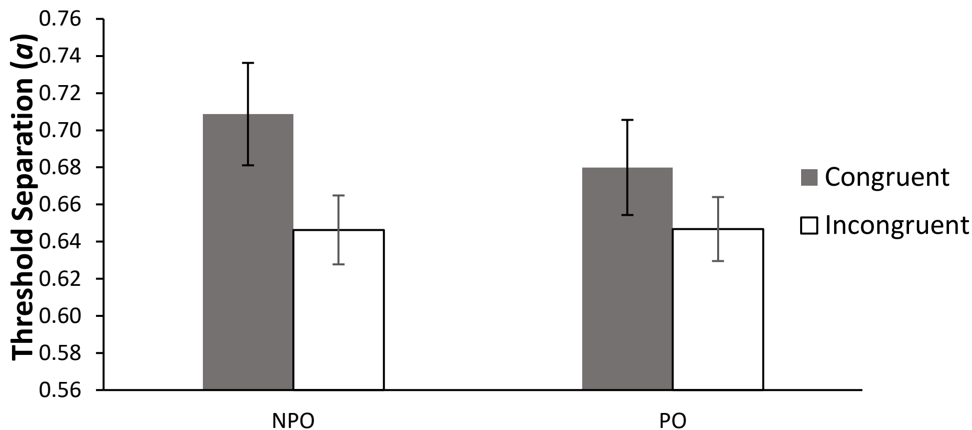

Figure 27. *Threshold separation for IAT trials (Congruent vs Incongruent) in both Social Conditions (NPO and PO). Errors bars denote one standard error around the mean.*

Nondecisions times (t0)

The analysis revealed a main effect of Block, *F* (1,93) =4.58, *p*=.04, η^2=.05, indicating that participants had slower nondecisions times in the Congruent block (*M*=414.31; *SD*=39.35) than in the Incongruent block (M=406.73; SD=39.62), see Figure 28.

The Social Condition main effect (*F*<1) and the Social Condition x Block interaction (*F*<1) were not significant.

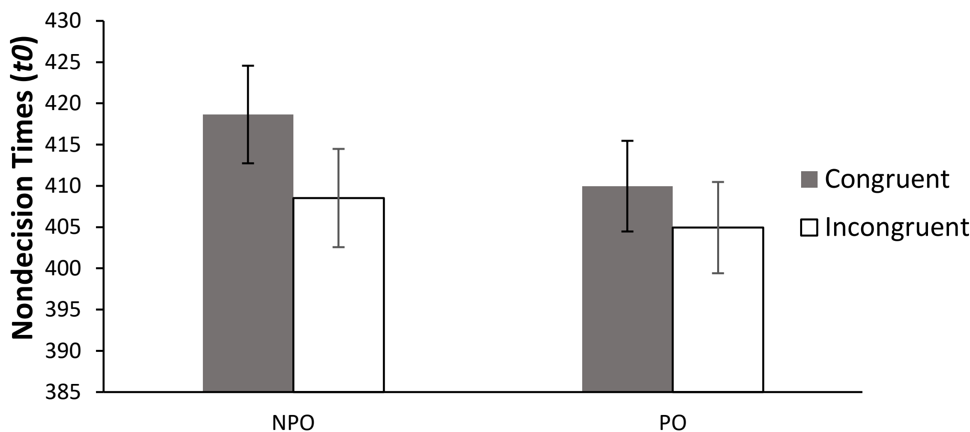

Figure 28. *Nondecision Times for IAT trials (Congruent vs Incongruent) in both Social Conditions (NPO and PO). Errors bars denote one standard error around the mean.*

Experiment 3- Discussion

In experiment 3, we modified the IAT used in experiment 2, adding more trials to the critical blocks and a response time restriction. Here we have replicated our findings reported on experiment 2, with participants in PO showing a higher stereotype bias (*d600*), than those in NPO. These results were again in the opposite direction of the effects described by Castelli and Tomelleri's (2008).

As in experiment 2, we conducted analysis targeting stereotype and SF effects in RTs, ERs and PDP components. Contrary to what we reported in experiment 2, we detected the stereotype effect in ERs and not in RTs, which may have been caused by the response time restriction. Moreover, we did not find effects of Social Condition across all measures, again raising our suspicion that *d600* as a measure of time and accuracy is more sensitive to SF effects.

The CDF analysis we performed found a stereotype bias for RT and ERs (contrary to experiment 2). Furthermore, the CDF analysis confirmed the pattern found in experiment 1 showing higher stereotype bias on the slowest responses (stereotype bias increases with RT). This effect was not moderated by Social Condition, although in the last bin the stereotype bias was lower in NPO, than in PO. This may suggest that when participants took longer to respond, it was easier to control the stereotype bias in the NPO condition.

The data analysis performed with the DM allowed us to detect stereotype effects in two parameters: *Drift-Rates* and *Relative Starting-Point*. These results suggest that participants gave more

correct responses (*Relative Starting-Point*) in the congruent block and were more efficient (*Drift-Rates*) handling stereotype information, than when facing counterstereotype information in the incongruent block.

We have only found marginal SF effects on *Drift-Rates*. This effect, although marginal, replicated the effects found with the *d600* index (experiment 2 and 3), indicating that the stereotype effect was stronger in PO than in NPO. Furthermore, these results support the conclusions of Lambert et al. (2003), suggesting more stereotyping in PO, and go against the findings of Castelli and Tomelleri (2008). Since participants in PO had more difficulty performing the IAT task, our results also match the SF theories that considered PO as a demanding context (e.g., Sander et al., 1978; Baron, 1986; Wagstaff et al., 2008).

In the DM analysis, we have also found unexpected results for *Threshold Separation* and *Nondecision Times*. According to Klauer et al. (2007), participants should have been more cautious and slower in their *Nondecision Times* for the Incongruent block, since the information presented was more difficult in this block. Our results, however, showed the opposite effect. Participants were more cautious and had slower *Nondecision Times* in the Congruent block. Although the IAT task we used (Race) was different from the IAT versions (Insect-Flower and Left-Right political view) in Klauer et al. (2007), it is more plausible to assume that these differences were due to our changes in the block presentation order. In IAT studies, the Incongruent block is usually presented before the Congruent block (see Greenwald et al., 2003), as in experiment 2. However, the presentation of the Incongruent block before the Congruent block could lead to a confound between the difficulty of handling counterstereotype information and task habituation effects (i.e., use of response keys), which was why we changed the block order and presented the Congruent before the Incongruent block. Our data suggests that block presentation order should be carefully planned in IAT studies, since having the congruent block first, lead to more caution and slower nonresponse times, probably because participants were still learning the correct match between their decisions and assigned response keys.

In conclusion, despite the changes made to the task (adding more trials and a response time restriction), the data from experiment 3 did not provide clear evidence on the role of PO on stereotyping. The results revealed the same pattern already found in experiment 2, simply suggesting a higher stereotype expression in PO, and that stereotype activation and influence on the participants' responses increases with response time.

Chapter V
General Discussion

This thesis has reviewed the SF literature and has offered an empirical approach to detecting SF effects on stereotyping. Since Triplett (1898), it has been proposed that PO influences human performance. The arrival of Cognitive Theories of SF (Bruning et al., 1968; Sanders & Baron, 1975; Sanders, et al., 1978; Baron et al., 1978; Baron, 1986) opened the possibility that PO also affects cognitive processes, an idea that has been strongly supported over the last two decades (e.g., Huguet et al., 1999; Muller & Butera, 2007; Wagstaff et al., 2008; Fonseca & Garcia-Marques, 2013). Although it is widely accepted that PO affects cognitive processes, few studies have been conducted to understand how SF works with more complex constructs such as stereotypes.

The present research strategically addressed this question by replicating the only two studies that have investigated how specific social conditions impact stereotyping. The first aim was to isolate the effects in a purer PO manipulation: coaction versus an alone condition. One of the two studies—by Lambert et al. (2003)—used a WIT paradigm, arguing that in social contexts individuals exert less cognitive control over their responses and hence display more stereotype bias. The other paper—by Castelli and Tomelleri (2008)—argued exactly the opposite using a Race IAT and an LDT: in PO, participants exhibit less stereotype bias owing to the activation of more egalitarian norms in PO. However, the two studies used different types of social conditions. The PO condition used in Lambert et al. (2003) was an accountability condition inside a booth, which can be confused with a real NPO condition. Castelli and Tomelleri (2008) compared co-action conditions with a condition that included the presence of the experimenter. By conducting the two tasks in actual SF conditions (PO as coaction and NPO without the presence of the experimenter or others), the present study sought to disentangle the apparent contradictions in these two results and to offer a better understanding of how SF occurs in stereotyping (i.e., stereotype bias). Thus, this is the first pure approach to determining SF effects on stereotyping.

While the aims were not fully realized, having conducted two sets of studies and more a complete analysis, it can be stated that it is not yet clear if and how SF effects occur in stereotyping. Following is a review of the main results to fill in details of this general conclusion.

Experiment 1 (WIT) did not offer evidence that stereotype errors (effect) are moderated by social conditions. These results oppose those of Lambert et al. (2003), who used the same WIT tasks but different social presences. The present results may suggest that either SF has no impact on stereotyping or PO and NPO manipulations should not be treated as equivalent to other Social Context manipulations (e.g., accountability; Lambert et al., 2003). However, the results were not null because there were unexpected (to the WIT) stereotype effects on RTs and on the PDP C *Component*. The PDP *C Component* evidenced different sensitivity to White and Black primes, and this pattern of results seemed to be detected only because of the addition of the NPO condition. Together, these results suggest that Social Conditions have an impact on stereotyping, although very subtle and without clear provenance.

Experiment 2 and 3 (with two different IAT versions) clearly showed that people in PO had higher *d600* than those in NPO. Therefore, SF in IATs causes stronger stereotype bias. However, these effects were opposite to those found by Castelli and Tomelleri (2008). Despite the *d600* being higher in PO than in NPO, nevertheless no stereotype effect was moderated by social condition when ERs were isolated. Rather, only a marginal moderation by social condition was detected in the pattern observed in RTs (experiment 2). Again, the more subtle manipulation of social presence seemed to have a more subtle impact on the evidence of stereotyping. This clearly suggests again that SF effects on stereotyping are subtle and not as straightforward as the reviewed papers suggest.

Even DM parameter analysis to WIT data does not seem to add much to this information; it merely suggests a higher *Threshold Separation* in PO than in NPO, and a marginal interaction between trials (Congruent vs Incongruent) and social conditions. This indicates that participants in PO spend more time in their decision processes and that only those in NPO, despite their faster decision processes, are more cautious during incongruent versus congruent trials. This opens the possibility that in WITs, social context does not affect stereotype activation but rather the way in which participants cope with the information they receive (an idea that is explored below).

The DM approach to the IAT data (time-restricted version) showed that SF effects in the IAT were marginally explained by the effect that PO exerted on *Drift-Rates*, indicating that participants in NPO can cope more efficiently with the IAT demands than those in PO. This corroborates the CDF analysis, which revealed that the stereotyping effect increases less across timed responses for those in NPO compared with those in PO. This pattern of results suggests that, relative to participants in NPO, those in PO impose less control over their responses, not more (as the literature review

suggested; see Huguet et al., 1999). As discussed below, PO may impose more cognitive demands on individuals, and thus cause more disturbance in their execution of control over the responses (Sanders & Baron, 1975; Sanders et al., 1978; Baron, 1986). Importantly, results do not suggest that by restricting attention, PO helps participants to better focus on the task goals (as Baron, 1986, assumed; see discussion below). An alternative reading (and more aligned with what was reviewed in Chapter I) is: rather than indicating less control in PO, the present results occurred because more information was activated in PO, and thus participants dealt with it relatively less efficiently.

From the summary above it is clear that no final conclusion can be made about the way that SF impacts stereotyping. SF effects on stereotyping were only detected in Experiments 2 and 3 (IAT) and not Experiment 1 (WIT). Importantly, when this research began, it was assumed that the divergent results of Lambert et al. (2003) and Castelli and Tomelleri (2008) were caused by differences in their experimental manipulations, and thus parsimonious results could be achieved by using pure SF effects (i.e., PO vs NPO). However, this was not the case; no direct SF effects were observed in the WIT (contradicting the results of Lambert et al., 2003), and more stereotyping was observed in PO than in NPO in IAT responses (contradicting the results of Castelli and Tomelleri, 2008).

While the present results did not completely answer the question of how SF affects stereotyping, they do offer insights about how SF affects cognition. Self-report measures obtained in Experiment 1 (WIT) indicated that those in PO felt the task was more demanding, as the Distraction-Conflict Theory (Sanders & Baron, 1975; Sanders et al., 1978; Baron et al., 1978) and the Overload Hypothesis (Baron, 1986) expect. PO contexts are perceived to be more distracting/overloading, and are thus understood as demanding more of the participants' capacity. Participants in PO were also faster and prone to more ERs and less control (PDP C *Component*) than those in NPO. This pattern indicates that participants in PO experience greater interference in the exerting of their executive control processes. However, it was not sufficient to promote clear SF effects on stereotyping.

On the other hand, the IAT data (Experiments 2 and 3) indicated that, compared with participants in NPO, those in PO stereotyped more and showed more evidence of stereotyping as their RTs increased (CDF). As referenced above, this may indicate that time increases the activation of information or that those in PO conditions have less capacity to deal with the information. As shown in the review, both of these possibilities are valid. Less control in PO can be explained by a greater activation of concepts in the mind (Allport, 1920; Fonseca & Garcia-Marques, 2013), which disrupts executive functions (Wagstaff et al., 2008). Stronger stereotype activation in PO may also be

caused by greater activation of stereotype concepts or by the use of dominant responses (Zajonc, 1965).

As stated above, the present data are unlikely to be explained either by a narrowing attention mechanism (Baron, 1986; Huguet et al., 1999) or by a response competition mechanism (Augustinova & Ferrand, 2012). Otherwise, better performance would be detected in PO, and thus more control. The present data also exclude the possibility that the observed effects were caused by evaluation effects (Cottrell et al., 1968), which disrupt performance, because the SF manipulations are not likely related to participants' self-report evaluations.

Why do these data not support the alternative explanations, that is, whether it is control that decreases in PO or whether it is stereotype activation that increases in PO? Because the tradeoff (activation vs control) was allowed, but was not perceived in DM and PDP measures as expected. One way to address this question in the future is to use overload manipulations orthogonally with social conditions (similar to Belletier & Camos, 2018). In this scenario, the condition (PO or NPO) that is less affected by overload would be deemed as already having scarce levels of control. Another way of addressing this question would be to use Devine's (1989) paradigm by asking participants about stereotype characteristics in PO and NPO, but adding the control of associative commonality (i.e., asking how the concepts are socially shared) as Fonseca and Garcia-Marques (2013) did in their Free Association Task. It would then be possible to map the spread activation of stereotype content in different SF conditions.

Because the conclusions made here are based on data from only two tasks (IAT and WIT), future studies should incorporate other stereotype tasks to find (or not) more parsimonious results. The literature indicates that performance in WITs and IATs are not correlated (Payne, 2005; Ito et al., 2015); thus, it is not surprising to have non-convergent evidence in both. It is very plausible that the two tasks capture stereotype effects in different ways because of their different characteristics. By using other types of tasks, it could be shown that the effect is sensitive to the different ways in which stereotyping is documented.

Nevertheless, an understanding of the differences between the IAT and the WIT can offer some insights into why it was easier to understand SF effects in stereotyping with the IAT task. For instance, the WIT is a sequential priming task that permits participants to ignore the prime (i.e., by focusing only on the target: Gun/Tool), and thereby ignore the stereotype interference. However, in IAT trials the categories (that work as primes) are always present simultaneously with the target

(Words/Faces), which does not allow participants to ignore the stereotype interference. Moreover, in WITs it is easier to focus on which information should be attended to (Target), since the decision is always between Tool and Gun. By contrast, in IATs, the demands vary between evaluating a word as positive or negative and evaluating a face as black or white. These extra demands imposed by different goals within the same task, known as switching costs (see Mierke & Klaeur, 2001), can also make anticipating—and thus ignoring—the stereotype interference more difficult.

The possibility of ignoring (or not) information that causes interference/biased responses is referred to in the literature as two different modes of control: early selection and later correction (see Schwarzkopp et al., 2016). Early selection control focuses on task goals. For instance, when discriminating between Tool and Gun in a WIT a participant can focus all their attention on the Target. Therefore, as Payne et al. (2005) verified with the WIT, early selection only captures errors when attention fails. The early selection control mechanism is also described in the literature as the C-First (Jacoby, 1991), the Discriminability (Conrey et al., 2005), or the Proactive approach (Braver, 2012). Later correction is a control mechanism that is engaged only when a participant detects conflict/interference/bias. In this way, the control works as a correction of bias information, and so is engaged in incongruent trials (e.g., Black-Tool WIT trials). At the same time, because this control depends on bias activation, correction failures lead to more bias influences in participants' responses (i.e., more congruency effects; Amodio & Swencionis, 2018). This latter correction mechanism is also known in the literature as A-First (Lindsay & Jacoby, 1994), Overcoming Bias (Conrey et al., 2005), or the Reactive approach (Braver, 2012).

The WIT seems to rely more on proactive processes as participants can correct their answers on a second attempt (Payne et al., 2005), that is, when they can correct their attentional failures. However, it does not happen in the same way in the FPST (Correl et al., 2015, experiment 3); participants replicate their errors on their second attempts, suggesting that their control failures make them more susceptible to bias/interference. From this reading of the literature, we can infer that the two tasks used in this thesis, WIT and IAT, likely rely on different forms of control. Specifically, a good performance in the WIT likely relies more on proactive control processes, and a good performance in the IAT likely depends more on the control of response interference and thus is highly dependent on reactive control (Conrey et al., 2005). These differences suggest that if SF modulates different types of control, it should impose performance differences in each of these different Tasks. Given that SF effects were more evident in IATs, reactive control was likely more efficiently exerted

by those in an NPO condition than those in a PO condition. By contrast, the required levels of proactive control over stereotype activation would be better balanced in both conditions (PO and NPO), since clear SF effects were not detected in the WIT. However, this interpretation conflicts with previous literature; Sharma et al. (2010), for example, demonstrated that in a Stroop task there was less interference (in that case, controlled by reactive control mechanisms) in PO than in NPO. Additionally, Augustinova and Ferrand (2012) observed less Stroop interference in PO than in NPO conditions, which for the authors was unrelated to a decision process. In other words, PO diminished response competition (i.e., interference that arises from the different response keys, not from the stimulus). Future research should directly address how the differences between the IAT and the WIT are related to varying efficiency in dealing with expected and actual interference.

In sum, the present data should be interpreted with caution, as SF effects on stereotyping seem to be more complex than initially thought, and different tasks may capture SF's influence on processing differently.

Further Questions

While the present data offer important information about SF effects on stereotyping, they also raise questions that should not be ignored in future research. Some of these questions are directly addressed below.

Why were some stereotype effects found in NPO?

Data collected from the WIT (experiment 1) for this thesis were the first to derive from a true NPO condition. Indeed, evidence of stereotyping was detected in both RTs and PDP *Control Components*. RT effects are not typically supposed to occur in the time-restricted (550 ms) WIT version (Payne et al., 2005), but should rather only appear in versions with unrestricted time (see Payne, 2001; Rivers, 2017). Therefore, these data indicate that social contexts can modulate how participants exert control over their responses.

Previous studies have suggested that the WIT relies on proactive control (and not on later correction processes), since there is no evidence of prime effects on the PDP *Control Components* nor on RTs in time-restricted versions (Payne, 2001; Payne et al., 2005), a pattern that was replicated in the present PO condition. However, the pattern detected in our NPO condition actually represents a later correction process (reactive), because there is evidence of interference correction (stereotype

effect on RTs and control differences in prime dependence). These different patterns in the data indicate that social context affected how participants coped with the task: those in PO had a proactive approach, and those in NPO a reactive approach. However, a proactive approach is linked to higher PDP *Control Component* (Amodio & Swencionis, 2018). Then, if PO promotes a proactive approach a higher PDP control component would be expected in PO; instead, a tendency for lower PDP control component was observed.

This contradiction challenges the present interpretation. Nevertheless, the literature does not provide any information about how proactive/reactive control works in different social contexts. It is possible that the lower PDP control observed in PO, compared with NPO, was attributable to social context and not to the use of different approaches. Future studies can address this question by introducing proactive/reactive control approaches (e.g., Braver, 2012; Amodio & Swencionis, 2018) while manipulating social context (NPO vs PO).

Thus, while the present data suggest that social context modulates participants' approaches in the WIT, many doubts remain about the type of mechanisms that are modulated.

What does the Diffusion Model say about these data?

The DM approach presents some evidence that PO impacts the DM parameters of each task differently: *Threshold Separation* on the WIT and *Drift-Rate* on the IAT. Specifically, our data for the WIT suggest that participants in NPO have higher *Threshold Separation* in incongruent trials than in congruent trials, a pattern that does not exist in PO. This pattern can be interpreted as evidence of a reactive approach, given that *Threshold Separation* is considered a cautious measure (e.g., Voss et al., 2013).

However, the present data could not reliably confirm that *Threshold Separation* was actually measuring a reactive approach. As observed, and contrary to Pleskac et al. (2017), threshold separation effects were not promoted by primes, but only by congruency (i.e., when white and black trials were conducted together). This may indicate that the effect is weak, and so can only be detected by including two congruent primes. This is a strange phenomenon, since stereotyping in the WIT should be stronger for Black-Tool trials (Payne, 2001), which would make the effect appear in the primes (Black vs White).

Another problem with the interpretation of the *Threshold Separation* index in experiment 1 is the fact that *Threshold Separation* was higher in PO than in NPO. This pattern does not make sense if

participants in PO made more errors and responded faster compared with participants in NPO; it represents an impulsive rather than a cautious approach.

One possibility for this incongruency may be that the DM had never been applied to WIT. In fact, the *Threshold Separation* has been understood as a measure of cautiousness where higher values represent searching for more information. However, in contrast to the FPST (Pleskac et al., 2017), the information (target) in the WIT is not on the screen when the participant responds (i.e., at the time of their decision). So, it is possible that in this scenario the measure represents a degree of uncertainty about the information, rather than a measure of cautiousness. If read this way, the data suggest that those in PO are generally more uncertain about their responses and those in NPO are more uncertain about their responses specifically in incongruent trials. This pattern makes even more sense if our previous results are accounted for (i.e., less PDP control component in PO, and less PDP control component for black trials in NPO).

Thus, the hypothesis that *Threshold Separation* is a measure of uncertainty instead of caution in WITs should be tested in future studies to fully understand the present findings. One way to do this is to measure individual self-reports of confidence and correlate them with the parameter. Another way is to use a WIT in which the target remains on the screen during the time of response; this would permit verification of a similar pattern to FPST data (Pleskac et al., 2017).

The DM in experiment 3 (IAT data) also detected a marginal effect in the *Drift-Rates*, which suggests that participants in PO were more affected by the stereotype information than those in NPO. Nevertheless, it remains unclear whether this effect is caused by less control and/or by more stereotype activation in PO. If the effect were stronger, differences in congruent and incongruent trials between social conditions could be compared. For instance, if higher *Drift-Rates* in congruent blocks were detected in PO rather than in NPO, it could be concluded that stereotype information is used (activated) more in PO. By contrast, if lower *Drift-Rates* in incongruent blocks were detected in PO rather than in NPO, it could be concluded that control processes engaged in IATs are less efficient in PO. However, as reported, the effect was weak, which did not permit this comparison between experimental cells. Thus, the DM approach to the present data was ineffective in providing an understanding of how SF affects the control-activation trade-off.

Despite these problems, the advantages of the DM approach in facilitating the incorporation of RT and accuracy data cannot be ignored. It allows us clearly state that stereotype effects are not related to *Relative Starting points* nor to *Nondecision Times*. Nevertheless, future research must explore

how DM parameters capture stereotype effects and how they behave in specific conditions (e.g., overload, egalitarian manipulations, and stereotype activation manipulations) to better understand their specific meaning in the stereotype literature.

Which Social Contexts reduce stereotyping?

The fact that SF effects are weak raises questions about their ecological and practical value. However, in some tasks, performance in PO is negatively affected, while in others it is positively affected (see Aiello & Douthitt, 2001). Additionally, even in a scenario in which PO is generally good for performance, some people are negatively affected because of individual differences (see Triplett, 1898; Allport, 1920; Uziel, 2007). Indeed, the application of SF findings to real-life problems must account for various features, such as the type of task and/or personality traits, to avoid decisions based only on experimental data. Aiello and Douthitt (2001) also discuss this reasoning; their concern is whether a performance that is boosted by working long periods of time in PO could impact stress levels. Therefore, the application of SF findings to real-life problems should be made with caution and should be followed up with ecologically valid studies.

Finally, the present data do not support a clear statement about how SF affects stereotyping. Thus, it is apparent that more must be done to reach ecological conclusions. For example, future studies should use ecological experiments such as: writing journal articles, which often use prejudiced terms (Ahdieh & Hahn, 1996); analyzing juridical cases, where stereotypical information influences the verdict and related punishment (see Graham & Lowery, 2004), and; CV analysis, where stereotype information also influences the hiring decision (e.g., Barbee & Gibson, 2001). Additionally, it must be noted that the studies on which this thesis is based focused only on the Black stereotype. While this focus was used in pursuit of resolving the inconsistent results of Lambert et al., (2003) versus Castelli and Tomelleri (2008), potential SF effects on stereotyping should also be tested with other stereotypes (e.g., woman and gipsy).

Do Social Facilitation effects really matter?

The greatest limitation of this thesis, and common in all mere presence studies, is the small effect sizes that SF exhibits, unlike other social effects such as evaluation effects (Bond & Titus, 1983). This is specifically problematic in this thesis because the experiments were designed on the basis of previous studies (Castelli & Tomelleri, 2008; Lambert et al., 2003) that did not use mere presence effects. Lambert et al. (2003) focused on accountability effects; while these are different from SF

effects, it was the only reference available to estimate the sample sizes for the present research. Furthermore, initially the SF manipulations used by Castelli and Tomelleri (2008) were adopted, but their NPO condition included the presence of a researcher, which can represent evaluation effects.

Evidently, small effect sizes and the use of inappropriate references to calculate those effects can produce relevant effects that are only marginal, and explain why some three-way interactions are not significant when simple analysis suggests differences. It is possible that with an appropriate sample size, more conclusive effects would appear.

So, if SF effects are small, why should we still study them? Although SF effects are small, the evidence for them does exist and can sometimes change the way that previous theories can be viewed. For instance, in experiment 1, the observed effects were only real in NPO (stereotype bias in RTs and prime differences in terms of PDP *Control Component*), which challenges the conclusion that WIT is a proactive task (Payne et al., 2005). Fonseca and Garcia-Marques (2013) also show that participants are more sensitive to the context in PO than in NPO, which indicates that social context really does change the way we think. Belletier et al. (2015), despite not seeing SF effects in the Simon task, found that the relation of working memory with the Simon interference was different in each Social Context.

Taken together, these findings teach us something that is universal in all experimental psychology: the lab setting matters. How much of the replicability crisis can be attributed to the varying methods and conditions of data collection across labs: individually, in a group, or in the presence of the experimenter? Even weak SF effects can be misguiding. How much do we really understand about psychology when we ignore the implications of the experimental setting? Are all theories supported in PO, or only in NPO? Because of these issues, the weak effects of SF cannot be ignored when they actually have strong consequences for the way in which the effects can be seen.

Conclusion

SF was initially studied only in terms of performance; now it is clear that a more comprehensive approach requires an investigation of the processes underlying performance. Such processes are present in many effects, including stereotype effects.

This thesis responded to this issue by replicating the experiments of two papers that had addressed how specific social conditions impact stereotyping (Lambert et al., 2003; Castelli & Tomelleri, 2008), with the addition of a more complete analysis of the effects. Unfortunately, the

effects that were detected did not replicate those of the previous studies. This shows that the use of actual SF conditions captures effects that are different from those detected in Lambert et al. (2003) and Castelli and Tomelleri (2008).

Despite this problem, the approach used in this thesis to reply to the initial question was efficient, and the findings offer guidance for new directions in the study of SF on stereotyping.

References

Ader, R., & Tatum, R. (1963). Individual and Paired Human Subjects '. *Journal of the Experimental Analysis of Behavior*, *6*(3), 357–359. https://doi.org/10.1901/jeab.1963.6-357

Agnew, N., & Agnew, M. (1963). Drive Level Effects on Tasks of Narrow and Broad Attention. *Quarterly Journal of Experimental Psychology*, *15*(1), 58–62. https://doi.org/10.1080/17470216308416552

Ahdieh, L., & Hahn, R. A. (1996). Use of the terms "race", "ethnicity", and "national origins": A review of articles in the American Journal of Public Health, 1980-1989. *Ethnicity and Health*, *1*(1), 95–98. https://doi.org/10.1080/13557858.1996.9961774

Aiello, J., & Douthitt, E. (2001). Social facilitation from Triplett to electronic performance monitoring. *Group Dynamics: Theory, Research, and Practice*, *5*(3), 163–180. https://doi.org/10.1037/1089-2699.5.3.163

Allport, F. H. (1920). The influence of the group upon association and thought. *Journal of Experimental Psychology,3*(3), 159-182. https://doi.org/10.1037/h0067891

Allport, F. (1924). Social psychology. New York: Houghton Mifflin.

Allport, G. W. (1954). *The nature of prejudice*. Oxford, England: Addison-Wesley.

Allport, G. W (1985). The historical background of modern social psychology. In Lindzey, G; Aronson, E (eds.). *The Handbook of Social Psychology* (pp.1-80). New York: McGraw Hill.

Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural Signals for the Detection of Unintentional Race Bias. *Psychological Science*, *15*(2), 88–93. https://doi.org/10.1111/j.0963-7214.2004.01502003.x

Amodio, D. M., & Swencionis, J. K. (2018). Proactive control of implicit bias: A theoretical model and implications for behavior change. *Journal of Personality and Social Psychology*, *115*(2), 255–275. https://doi.org/10.1037/pspi0000128

Anderson-Hanley, C., Snyder, A. L., Nimon, J. P., & Arciero, P. J. (2011). Social facilitation in virtual reality-enhanced exercise: Competitiveness moderates exercise effort of older adults. *Clinical*

Interventions in Aging, *6*(1), 275–280. https://doi.org/10.2147/CIA.S25337

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261–295. https://doi.org/10.1016/S0022-5371(83)90201-3

Asch, S.E. (1951). Effects of group pressure on the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership and men*(pp. 177–190). Pittsburgh, PA:Carnegie Press.

Augustinova, M., & Ferrand, L. (2012). The influence of mere social presence on Stroop interference: New evidence from the semantically-based Stroop task. *Journal of Experimental Social Psychology*, *48*, 1213–1216. https://doi.org/10.1016/j.jesp.2012.04.014

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), The psychology of learning and motivation Vol. 8 (pp. 47–89). New York: Academic Press.

Barbee, E., & Gibson, S. (2001). Our dismal progress: The recruitment of non-whites into nursing. *Journal of Nursing Education, 40*, 243-244. https://doi.org/10.3928/0148-4834-20010901-03

Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 3-51). New York: Guilford Press

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype-activation on action. *Journal of Personality and Social Psychology*, *71*(2), 230–244. https://doi.org/ 10.1037//0022-3514.71.2.230

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The Generality of the Automatic Attitude Activation Effect. *Journal of Personality and Social Psychology*, *62*(6), 893–912. https://doi.org/10.1037/0022-3514.62.6.893

Baron, R. S. (1986). Distraction-conflict theory: Progress and problems. *Advances in Experimental Social Psychology*, *19*(C), 1–40. https://doi.org/10.1016/S0065-2601(08)60211-7

Baron, R. S., Moore, D., & Sanders, G. S. (1978). Distraction as a source of drive in social facilitation research. *Journal of Personality and Social Psychology, 36*(8), 816–824. https://doi.org/10.1037/0022-3514.36.8.816

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process

tree modeling. *Psychonomic Bulletin and Review*, *6*(1), 57–86. https://doi.org/10.3758/BF03210812

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin*, *91*(1), 3–26. https://doi.org/10.1037/0033-2909.91.1.3

Belletier, C., & Camos, V. (2018). Does the experimenter presence affect working memory? *Annals of the New York Academy of Sciences*, *1424*(1), 212–220. https://doi.org/10.1111/nyas.13627

Belletier, C., Davranche, K., Tellier, I. S., Dumas, F., Vidal, F., Hasbroucq, T., & Huguet, P. (2015). Choking under monitoring pressure: being watched by the experimenter reduces executive attention. *Psychonomic Bulletin and Review*, *22*(5), 1410–1416. https://doi.org/10.3758/s13423-015-0804-9

Bergum, B. O., & Lehr, D. J. (1963). Effects of authoritarianism on vigilance performance. *Journal of Applied Psychology*, *47*(1), 75–77. https://doi.org/10.1037/h0043188

Berns, G. S., Cohen, J. D., & Mintun, M. A. (1997). Brain regions responsive to novelty in the absence of awareness. *Science*, *276*(5316), 1272–1275. https://doi.org/10.1126/science.276.5316.1272

Bodenhausen, G. V., Todd, A. R., & Becker, A. P. (2006). Categorizing the Social World: Affect, Motivation, and Self-Regulation. In A. Markman, & R. Brian (Eds.), *Categories in use* (pp. 123-155). https://doi.org/10.1016/S0079-7421(06)47004-3

Bond, C. F. (1982). Social facilitation: A self-presentational view. *Journal of Personality and Social Psychology*, *42*(6), 1042–1050. https://doi.org/10.1037/0022-3514.42.6.1042

Bond, C F, & Titus, L. J. (1983). Social facilitation: a meta-analysis of 241 studies. *Psychological Bulletin*, *94*(2), 265–292. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6356198

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in Cognitive Sciences*, *16*(2), 106–113. https://doi.org/10.1016/j.tics.2011.12.010

Brewer, M. B. (1988). A dual process model of impression formation. In T.K. Krull & R.S. Wyer (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc

Bruning, J. L., Capage, Ja. E., Kozuh, G. F., Young, P. F., & Young, W. E. (1968). Socially Induced Drive and Range of Cue Utilization. *Journal of Personality and Social Psychology*, *9*(3), 242–244. https://doi.org/10.1037/h0025895

Bub, D. N., Masson, M. E. J., & Lalonde, C. E. (2006). Cognitive control in children: Stroop interference and suppression of word reading. *Psychological Science*, *17*(4), 351–357. https://doi.org/10.1111/j.1467-9280.2006.01710.x

Burle, B., Spieser, L., Servant, M., & Hasbroucq, T. (2014). Distributional reaction time properties in the Eriksen task: marked differences or hidden similarities with the Simon task? *Psychonomic Bulletin & Review*, *21*(4), 1003–1010. https://doi.org/10.3758/s13423-013-0561-6

Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring Psychological Significance from Physiological Signals. *American Psychologist*, *45*(1), 16–28. https://doi.org/10.1037/0003-066X.45.1.16

Carver, C. S., & Scheier, M. F. (1981). The self-attention-induced feedback loop and social facilitation. *Journal of Experimental Social Psychology*, *17*(6), 545–568. https://doi.org/10.1016/0022-1031(81)90039-1

Castelli, L., & Tomelleri, S. (2008). Contextual effects on prejudiced attitudes: When the presence of others leads to more egalitarian responses. *Journal of Experimental Social Psychology*, *44*(3), 679–686. https://doi.org/10.1016/j.jesp.2007.04.006

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology*, *33*(5), 541–560. https://doi.org/10.1006/jesp.1997.1329

Cohen, S. (1978). Environmental load and the allocation of attention. In A. Baum, J. E. Singer, & S. Valins (Eds.), *Advances in environmental psychology* (pp. 1-30). Hillsdale, NJ: Erlbaum.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407

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

Correll, J., Crawford, M., Wiitenbrink, B., & Sadler, M. S. (2015). Stereotypic Vision : How Stereotypes Disambiguate Visual Stimuli. *Journal of Personality and Social Psychology*, *108*(2), 219–233. https://doi.org/ 10.1037/pspa0000015

Cottrell, N. B. (1972). Social facilitation. In C. G. McClintock (Ed.), *Experimental social psychology* (pp. 185–236). New York: Holt, Rinehart and Winston.

Cottrell, N. B., Wack, D. L., Sekerak, G. J., & Rittle, R. H. (1968). Social facilitation of dominant responses by the presence of an audience and the mere presence of others. *Journal of Personality and Social Psychology*, *9*(3), 245–250. https://doi.org/10.1037/h0025902

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Dashiell, J. F. (1930). An experimental analysis of some group effects. *The Journal of Abnormal and Social Psychology*, *25*(2), 190–199. https://doi.org/10.1037/h0075144

Davranche, K., Hall, B., & McMorris, T. (2009). Effect of acute exercise on cognitive control required during an Eriksen flanker task. *Journal of Sport & Exercise Psychology*, *31*(5), 628–639. https://doi.org/10.1123/jsep.31.5.628

Davranche, K., & McMorris, T. (2009). Specific effects of acute moderate exercise on cognitive control. *Brain and Cognition*, *69*(3), 565–570. https://doi.org/10.1016/j.bandc.2008.12.001

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

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

Devine, P. G., & Elliot, A. J. (1995). Are Racial Stereotypes Really Fading? The Priceton Trilogy Revisited. *Personality and Social Psychology Bulletin*, *21*(11), 1139–1150. https://doi.org/https://doi.org/10.1177/01461672952111002

Doherty, M. J., Tsuji, H., & Phillips, W. A. (2008). The context sensitivity of visual size perception varies across cultures. *Perception*, *37*(9), 1426–1433. https://doi.org/10.1068/p5946

Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of Blacks. *Journal of Personality and Social Psychology*, *34*(4), 590–598. https://doi.org/10.1037/0022-3514.34.4.590

Duval, S., & Wicklund, R. A. (1972). *A theory of objective self-awareness*. Academic Press.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*(3), 183–201. https://doi.org/10.1037/h0047707

Engle, R. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23. Retrieved from http://cdp.sagepub.com/content/11/1/19.short

Esses, V. M., Haddock, G., & Zanna, M.P. (1993) Values, Stereotypes, and Emotions as Determinants of Intergroup Attitudes. In *Affect, Cognition, and Stereotyping: Interactive Processes in Group Perception*, edited by Diane M. Mackie, and David Lewis Hamilton, 137–166. San Diego: Academic Press.

Fazio, R.H. (1986) How Do Attitudes Guide Behavior? In: Sorrentino, R.M. and Higgins, E.T., Eds., *The Handbook of Motivation and Cognition: Foundation of Social Behavior*, Guilford Press, New York, 204-243.

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

Figueira, P., & Garcia-Marques, T. (2019). The other side of self-monitoring: Inhibition control in and out a social context. *Análise Psicológica*, *37*(1), 29–39. https://doi.org/10.14417/ap.1498

Fiske, S. T., & Neuberg, S. L. (1990). A Continuum of Impression Formation, from Category-Based to Individuating Processes: Influences of Information and Motivation on Attention and Interpretation. *Advances in Experimental Social Psychology*, *23*(C), 1–74. https://doi.org/10.1016/S0065-2601(08)60317-2

Fiske, S. T., & Taylor, S. E. (2013). *Social cognition: From brains to culture*. Atlanta, GA: SAGE.

Fonseca, R., & Garcia-Marques, T. (2013). Back to Basics: Socially Facilitated Situated Cognition. *Social Cognition*, *31*(2), 147–161. https://doi.org/10.1521/soco.2013.31.2.147

Freeman, J. B., & Ambady, N. (2011). A Dynamic Interactive Theory of Person Construal. *Psychological Review*, *118*(2), 247–279. https://doi.org/10.1037/a0022327

Garcia-Marques, T., Fernandes, A., Fonseca, R., & Prada, M. (2015a). Social presence and the composite face effect. *Acta Psychologica*, *158*(May), 61–66. https://doi.org/10.1016/j.actpsy.2015.04.001

Garcia-Marques, T., Fernandes, A., Prada, M., Fonseca, R., & Hagá, S. (2015b). Seeing the Big Picture: Size Perception Is More Context Sensitive in the Presence of Others. *Plos One*, *10*(11), e0141992. https://doi.org/10.1371/journal.pone.0141992

Gawronski, B. (2009). Ten frequently asked questions about implicit measures and their frequently supposed, but not entirely correct answers. *Canadian Psychology*, *50*(3), 141–150. https://doi.org/10.1037/a0013848

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60*(4), 509–517. https://doi.org/10.1037/0022-3514.60.4.509

Goffman, E. (*1959*). *The presentation of self in everyday life*. Garden City: Doubleday.

Govorun, O., & Payne, B. K. (2006). Ego – Depletion and Prejudice : Separating Automatic and Controlled Components. *Social Cognition*, *24*(2), 111–136. https://doi.org/ 10.1521/soco.2006.24.2.111

Graham, S., & Lowery, B. S. (2004). Priming unconscious racial stereotypes about adolescent offenders. *Law and Human Behavior*, *28*(5), 483–504. https://doi.org/10.1023/B:LAHU.0000046430.65485.1f

Grant, T., & Dajee, K. (2003). Types of task, types of audience, types of actor: interactions between mere presence and personality type in a simple mathematical task. *Personality and Individual Differences*, *35*(3), 633–639. https://doi.org/10.1016/S0191-8869(02)00241-6

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

1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Greenwald, Anthony G, & Nosek, B. A. (2003). Understanding and Using the Implicit Association Test : I . An Improved Scoring Algorithm, *85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Guerin, B. (1993). *Social facilitation*. Cambridge, England: Cambridge University Press.

Guerin, B., & Innes, J. (1984). Explanations of social facilitation: A review. *Current Psychology Research & Reviews*, *3*(2), 32–52. Retrieved from http://link.springer.com/article/10.1007/BF02686548

Hall, B., & Henningsen, D. D. (2008). Social facilitation and human–computer interaction. *Computers in Human Behavior*, *24*(6), 2965–2971. https://doi.org/10.1016/j.chb.2008.05.003

Hamilton, D. L., and Sherman, J. W. (1994). "Stereotypes," in *Handbook of Social Cognition*, *Applications*, eds R. S. Wyer Jr. and T. K. Srull (Hillsdale, NJ: Erlbaum), 1–68.

Harkins, S. G. (1987). Social loafing and social facilitation. *Journal of Experimental Social Psychology*, *23*(1), 1–18. https://doi.org/10.1016/0022-1031(87)90022-9

Henchy, T., & Glass, D. C. (1968). Evaluation apprehension and the social facilitation of dominant and subordinate responses. *Journal of Personality and Social Psychology*, *10*(4), 446–454. https://doi.org/10.1037/h0026814

Henry, J. D., & Crawford, J. R. (2004). A meta-analytic review of verbal fluency performance in patients with traumatic brain injury. *Neuropsychology*, *18*(4), 621–628. https://doi.org/10.1037/0894-4105.18.4.621

Hirshorn, E. A., & Thompson-Schill, S. L. (2006). Role of the left inferior frontal gyrus in covert word retrieval: Neural correlates of switching during verbal fluency. *Neuropsychologia*, *44*(12), 2547–2557. https://doi.org/10.1016/j.neuropsychologia.2006.03.035

Hobson, N. M., & Inzlicht, M. (2016). The mere presence of an outgroup member disrupts the brain's feedback-monitoring system. *Social Cognitive and Affective Neuroscience*, *11*(11), 1698–1706. https://doi.org/10.1093/scan/nsw082

Houghton, G., & Grange, J. a. (2011). CDF-XL: computing cumulative distribution functions of reaction time data in Excel. *Behavior Research Methods*, *43*(4), 1023–1032.

https://doi.org/10.3758/s13428-011-0119-3

Huguet, P., Galvaing, M. P., Monteil, J. M., & Dumas, F. (1999). Social presence effects in the Stroop task: Further evidence for an attentional view of social facilitation. *Journal of Personality and Social Psychology*, *77*(5), 1011–1024. https://doi.org/10.1037/0022-3514.77.5.1011

Husband, R. W. (1931). Analysis of Methods in Human Maze Learning. *The Pedagogical Seminary and Journal of Genetic Psychology*, *39*(2), 258–278. https://doi.org/10.1080/08856559.1931.10532308

Isenberg,D. J. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology*, *50*(6), 1141–1151. https://doi.org/10.1037/0022-3514.50.6.1141

Ito, T. A., Friedman, N. P., Bartholow, B. D., Correll, J., Loersch, C., Altamirano, L. J., & Miyake, A. (2015). Toward a comprehensive understanding of executive cognitive function in implicit racial bias. *Journal of Personality and Social Psychology*, *108*(2), 187–218. https://doi.org/10.1037/a0038557

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

Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating Conscious and Unconscious Influences of Memory: Measuring Recollection. *Journal of Experimental Psychology: General*, *122*(2), 139–154. https://doi.org/10.1037/0096-3445.122.2.139

Janis, I. (1991). Groupthink. In E. Griffin (Ed.) *A First Look at Communication Theory* (pp. 235 - 246). New York: McGrawHill.

Johnson, D. J., Hopwood, C. J., Cesario, J., & Pleskac, T. J. (2017). Advancing Research on Cognitive Processes in Social and Personality Psychology: A Hierarchical Drift Diffusion Model Primer. *Social Psychological and Personality Science*, *8*(4), 413–423. https://doi.org/10.1177/1948550617703174

Jonides, J. (1980). Towards a model of the Mind's Eye's movement. *Canadian Journal of Psychology/Revue canadienne de psychologie, 34*(2), 103–112. https://doi.org/10.1037/h0081031

Judd, C. M., Blair, I. V., & Chapleau, K. M. (2004). Automatic stereotypes vs. automatic prejudice: Sorting out the possibilities in the Payne (2001) weapon paradigm. *Journal of Experimental Social*

Psychology, *40*(1), 75–81. https://doi.org/10.1016/S0022-1031(03)00063-5

Katz, D., & Braly, K. (1933). Racial stereotypes of one hundred college students. *Journal of Abnormal and Social Psychology*, *28*(3), 280–290. https://doi.org/10.1037/h0074049

Kitayama, S., Duffy, S., Kawamura, T., & Larsen, J. T. (2003). Perceiving an Object and Its Context in Different Cultures: A Cultural Look at New Look. *Psychological Science*, *14*(3), 201–206. https://doi.org/10.1111/1467-9280.02432

Klauer, K. C., Herfordt, J., & Voss, A. (2008). Social presence effects on the stroop task: Boundary conditions and an alternative account. *Journal of Experimental Social Psychology*, *44*(2), 469–476. https://doi.org/10.1016/j.jesp.2007.02.009

Klauer, K. C., & Voss, A. (2008). Effects of race on responses and response latencies in the weapon identification task: A test of six models. *Personality and Social Psychology Bulletin*, *34*(8), 1124–1140. https://doi.org/10.1177/0146167208318603

Klauer, K. C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of the Implicit Association Test: A diffusion-model analysis. *Journal of Personality and Social Psychology*, *93*(3), 353–368. https://doi.org/10.1037/0022-3514.93.3.353

Kunda, Z. (1999) *Social cognition: Making sense of people*. Cambridge, MA: The MIT Press.

Kunda, Z., & Thagard, P. (1996). Forming Impressions from Stereotypes, Traits, and Behaviors: A Parallel-Constraint-Satisfaction Theory. *Psychological Review, 103*(2), 284-308. https://doi.org/10.1037/0033-295X.103.2.284

Lambert, A. J., Payne, B. K., Jacoby, L. L., Shaffer, L. M., Chasteen, A. L., & Khan, S. R. (2003). Stereotypes as dominant responses: On the "social facilitation" of prejudice in anticipated public contexts. *Journal of Personality and Social Psychology*, *84*(2), 277–295. https://doi.org/10.1037/0022-3514.84.2.277

Landers, D. M., & McCullagh, P. D. (1976). Social facilitation of motor performance. *Exercise and Sport Sciences Reviews*, *4*(1), 125–162. https://doi.org/10.1249/00003677-197600040-00006

Lasagna, L., & McCann, W. P. (1957). Effect of "tranquilizing" drugs on amphetamine toxicity in aggregated mice. *Science*, *125*(3260), 1241–1242. https://doi.org/10.1126/science.125.3260.1241

Leary, M. R., & Baumeister, R. F. (2000). The nature and function of self-esteem: Sociometer theory. *Advances in Experimental Social Psychology*, *32*, 1–62. https://doi.org/10.1016/s0065-2601(00)80003-9

Lepore, L., & Brown, R. (1997). Category and Stereotype Activation: Is Prejudice Inevitable? *Journal of Personality and Social Psychology*, *72*(2), 275–287. https://doi.org/10.1037/0022-3514.72.2.275

Lerche, V., Voss, A., & Nagler, M. (2017). How many trials are required for parameter estimation in diffusion modeling? A comparison of different optimization criteria. *Behavior Research Methods*, *49*(2), 513–537. https://doi.org/10.3758/s13428-016-0740-2

Lindsay, D. S., & Jacoby, L. L. (1994). Stroop process dissociations: the relationship between facilitation and interference. *Journal of Experimental Psychology. Human Perception and Performance*, *20*(2), 219–234. https://doi.org/ 10.1037//0096-1523.20.2.219

Lippmann, W. (1922). Stereotypes. In W. Lippmann, *Public opinion* (p. 79–94). MacMillan Co.

Maass, A., & Clark, R. D. (1983). Internalization versus compliance: Differential processes underlying minority influence and conformity. *European Journal of Social Psychology*, *13*(3), 197–215. https://doi.org/10.1002/ejsp.2420130302

Maass, A., & Clark, R. D. (1986). Conversion theory and simultaneous majority/minority influence: Can reactance offer an alternative explanation? *European Journal of Social Psychology*, *16*(3), 305–309. https://doi.org/10.1002/ejsp.2420160307

Macrae, C. N., Stangor, C., & Milne, A. B. (1994). Activating social stereotypes: A functional analysis. *Journal of Experimental Social Psychology*. https://doi.org/10.1006/jesp.1994.1018

Macrae, C. N., & Bodenhausen, G. V. (2000). Social Cognition: Thinking Categorically about Others. *Annual Review of Psychology*, *51*(1), 93–120. https://doi.org/10.1146/annurev.psych.51.1.93

Macrae, C. N., Hewstone, M., & Griffiths, R. J. (1993). Processing load and memory for stereotype-based information. *European Journal of Social Psychology*, *23*, 77–87. https://doi.org/10.1002/ejsp.2420230107

Macrae, C. N., Bodenhausen, G. V, & Milne, A. B. (1995). The Dissection of Selection in Person Perception: Inhibitory Processes in Social Stereotyping. *Journal of Personality and Social Psychology*

69(3), 397–407. https://doi.org/ 10.1037//0022-3514.69.3.397

Macrae, C. N., Bodenhausen, G. V, Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology*, *67*(5), 808–817. https://doi.org/10.1037/0022-3514.67.5.808

Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices - a peek inside the cognitive toolbox. *Journal of Personality and Social Psychology*, *66*(1), 37–47. https://doi.org/10.1037/0022-3514.66.1.37

Maddox, K. B., & Chase, S. G. (2004). Manipulating subcategory salience: Exploring the link between skin tone and social perception of Blacks. *European Journal of Social Psychology*, *34*(5), 533–546. https://doi.org/10.1002/ejsp.214

Markus, H. (1978). The Effect of Mere Presence on Social: An Unobtrusive Test. *Journal of Experimental Soci*, *14*, 389–397. https://doi.org/10.1016/0022-1031(78)90034-3

Mason, J. W., & Brady, J. V. (1956). Plasma 17-Hydroxycorticosteroid Changes Related to Reserpine Effects on Emotional Behavior. *Science*, *124*(3229), 983–984. https://doi.org/10.1126/science.124.3229.983

McConahay J.B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J.F. Dovidio and S.L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91–125). Orlando, FL: Academic Press.

Mierke, J., & Klauer, K. C. (2001). Implicit association measurement with the IAT: Evidence for effects of executive control porcesses. *Zeitschrift für Experimentelle Psychologie, 48*(2), 107–122. https://doi.org/10.1026//0949-3946.48.2.107

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Moscovici, S., & Personnaz, B. (1980). Studies in social influence. V. Minority influence and conversion behavior in a perceptual task. *Journal of Experimental Social Psychology*, *16*(3), 270–282. https://doi.org/10.1016/0022-1031(80)90070-0

Moscovic, S., & Personnaz, B. (1986). Studies on latent influence by the spectrometer method I: The impact of psychologization in the case of conversion by a minority or a majority. *European Journal of Social Psychology*, *16*(4), 345–360. https://doi.org/10.1002/ejsp.2420160404

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology*, *77*(1), 167–184. https://doi.org/10.1037/0022-3514.77.1.167

Moskowitz, G. B., & Li, P. (2011). Egalitarian goals trigger stereotype inhibition: A proactive form of stereotype control. *Journal of Experimental Social Psychology*, *47*(1), 103–116. https://doi.org/10.1016/j.jesp.2010.08.014

Moskowitz, G. B., Salomon, A. R., & Taylor, C. M. (2000). Preconsciously Controlling Stereotyping: Implicitly Activated Egalitarian Goals Prevent the Activation of Stereotypes. *Social Cognition*, *18*(2), 151–177. https://doi.org/10.1521/soco.2000.18.2.151

Muller, D., Atzeni, T., & Butera, F. (2004). Coaction and upward social comparison reduce the illusory conjunction effect: Support for distraction-conflict theory. *Journal of Experimental Social Psychology*, *40*(5), 659–665. https://doi.org/10.1016/j.jesp.2003.12.003

Muller, D., & Butera, F. (2007). The Focusing Effect of Self-Evaluation Threat in Coaction and Social Comparison. *Journal of Personality and Social Psychology*, *93*(2), 194–211. https://doi.org/10.1037/0022-3514.93.2.194

Müller, H. J., & Rabbitt, P. M. A. (1989). Reflexive and Voluntary Orienting of Visual Attention: Time Course of Activation and Resistance to Interruption. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(2), 315–330. https://doi.org/10.1037/0096-1523.15.2.315

Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin*, *83*(4), 602–627. https://doi.org/10.1037/0033-2909.83.4.602

Na, J., & Kitayama, S. (2012). Will people work hard on a task they choose? Social-eyes priming in different cultural contexts. *Journal of Experimental Social Psychology*, *48*(1), 284–290.

https://doi.org/10.1016/j.jesp.2011.09.003

Neely, J. H., & Kahan, T. A. (2001). Is semantic activation automatic? A critical re-evaluation. In H. L. Roediger III, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *Science conference series. The nature of remembering: Essays in honor of Robert G. Crowder* (p. 69–93). American Psychological Association. https://doi.org/10.1037/10394-005

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P. H., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology, 38*(5), 752–760. https://doi.org/10.1111/1469-8986.3850752

Ozonoff, S., & Jensen, J. (1999). Brief Report: Specific Executive Function Profiles in Three Neurodevelopmental Disorders. *Journal of Autism and Developmental Disorders, 29***,** 171-177. https://doi.org/10.1023/A:1023052913110

Park, S., & Catrambone, R. (2007). Social facilitation effects of virtual humans. *Human Factors*, *49*(6), 1054–1060. https://doi.org/10.1518/001872007X249910

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, *81*(2), 181–192. https://doi.org/10.1037/0022-3514.81.2.181

Payne, B K. (2005). Conceptualizing control in social cognition: how executive functioning modulates the expression of automatic stereotyping. *Journal of Personality and Social Psychology*, *89*(4), 488–503. https://doi.org/10.1037/0022-3514.89.4.488

Payne, B.K. (2006). Weapon bias: split-second decisions and unintended stereotyping. *Current Directions in Psychological Science*, *15*(6), 287–291. https://doi.org/10.1111/j.1467-8721.2006.00454.x

Payne, B. K., Shimizu, Y., & Jacoby, L. L. (2005). Mental control and visual illusions: Toward explaining race-biased weapon misidentifications. *Journal of Experimental Social Psychology*, *41*(1), 36–47. https://doi.org/10.1016/j.jesp.2004.05.001

Pessin, J. (1933). The comparative effects of social and mechanical stimulation on memorizing. *American Journal of Psychology*, *45*(Journal Article), 263–270.

Pessin, J., & Husband, E. W. (1933). Effects of social stimulation on human maze learning. *The*

Journal of Abnormal and Social Psychology, *28*(2), 148–154. https://doi.org/10.1037/h0074711

Phillips, L. H., Bull, R., Adams, E., & Fraser, L. (2002). Positive Mood and Executive Function. Evidence From Stroop and Fluency Tasks. *Emotion*, *2*(1), 12–22. https://doi.org/10.1037/1528-3542.2.1.12

Pleskac, T. J., Cesario, J., & Johnson, D. J. (2017). How race affects evidence accumulation during the decision to shoot. *Psychonomic Bulletin and Review*, 1–30. https://doi.org/10.3758/s13423-017-1369-6

Posner M.I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Lawrence Erlbaum.

Posner, M. I. (1980). Orienting of attention. *Journal of Experimental Psychology*, *32*(July 1979), 3–25. https://doi.org/1 0.1080/00335558008248231

Posner, M.I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D.G. Bouwhuis (Eds.), *Attention and performance X. Control of language processes.* Hillsdale, NJ: Erlbaum. Psychology

Software Tools, Inc. [E-Prime 3.0]. (2016). Retrieved from https://www.pstnet.com.

Puntoni, S., & Tavassoli, N. (2007). Social context and advertising memory. *Journal of Marketing Research*, *XLIV*(May), 284–296. Retrieved from http://www.journals.marketingpower.com/doi/abs/10.1509/jmkr.44.2.284

Rajecki, D., Ickes, W., Corcoran, C., & Lenerz, K. (1977). Social facilitation of human performance: Mere presence effects. *The Journal of Social Psychology, 102*(2), 297–310. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/00224545.1977.9713277

Rakić, T., Steffens, M. C., & Mummendey, A. (2011). When it matters how you pronounce it: The influence of regional accents on job interview outcome. *British Journal of Psychology*, *102*(4), 868–883. https://doi.org/10.1111/j.2044-8295.2011.02051.x

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, *86*(3), 446–461. https://doi.org/10.1037/0033-2909.86.3.446

Ratcliff, R., & Rouder, J. N. (1998). Modeling Response Times for Two-Choice Decisions.

Psychological Science, *9*(5), 347–356. https://doi.org/10.1111/1467-9280.00067

Ratcliff, R., Thapar, A., & McKoon, G. (2004). A diffusion model analysis of the effects of aging on recognition memory. *Journal of Memory and Language*, *50*(4), 408–424. https://doi.org/10.1016/j.jml.2003.11.002

Remington, R. W., Johnston, J. C., & Yantis, S. (1992). Involuntary attentional capture by abrupt onsets. *Perception & Psychophysics*, *51*(3), 279–290. https://doi.org/10.3758/BF03212254

Rende, B., Ramsberger, G., & Miyake, A. (2002). Commonalities and differences in the working memory components underlying letter and category fluency tasks: A dual-task investigation. *Neuropsychology*, *16*(3), 309–321. https://doi.org/10.1037/0894-4105.16.3.309

Ridderinkhof, K. R. (2002). Activation and suppression in conflict tasks: Empirical clarification through distributional analyses. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action. Attention and performance* (Vol. 19, pp. 494–519). Oxford: Oxford University Press

Ridderinkhof, K. R., Scheres, A., Oosterlaan, J., & Sergeant, J. a. (2005). Delta plots in the study of individual differences: new tools reveal response inhibition deficits in AD/Hd that are eliminated by methylphenidate treatment. *Journal of Abnormal Psychology*, *114*(2), 197–215. https://doi.org/10.1037/0021-843X.114.2.197

Ridderinkhof, K. R., van den Wildenberg, W. P. M., Wijnen, J., & Burle, B. (2004). Response Inhibition in Conflict Tasks Is Revealed in Delta Plots. In M. I. Posner (Ed.), *Cognitive neuroscience of attention* (p. 369–377). The Guilford Press.

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial Modeling and the Measurement of Cognitive Processes. *Psychological Review*, *95*(3), 318–339. https://doi.org/10.1037/0033-295X.95.3.318

Ringelmann, M. (1913). Appareils de cultur mecanique avec treuils et cables (resultats d'essais) [Mechanical tilling equipment with winches and cables (results of tests)]. Annales de I'lnstitut National Agronomique, 2e serie—tome XII, 299-343.

Rivers, A. M. (2017). The Weapons Identification Task : Recommendations for adequately powered research. *Plos One*, *12*(6), 1–15.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology General*, *126*(3), 211–227. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9281831

Sanders, G. S. (1981). Driven by Distraction : An Integrative Review of Social Facilitation Theory and Research. *Journal of Experimental Social Psychology*, *17*, 227–251. https://doi.org/10.1016/0022-1031(81)90024-X

Sanders, G. S., & Baron, R. S. (1975). The Motivating Effects of Distraction on Task Performance. *Journal of Personality and Social Psychology*, *32*(6), 956–963. https://doi.org/10.1037/0022-3514.32.6.956

Sanders, G. S., Baron, R. S., & Moore, D. L. (1978). Distraction and social comparison as mediators of social facilitation effects. *Journal of Experimental Social Psychology*, *14*(3), 291–303. https://doi.org/10.1016/0022-1031(78)90017-3

Schmidt, J. R., & Cheesman, J. (2005). Dissociating Stimulus-Stimulus and Response-Response Effects in the Stroop Task. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *59*(2), 132–138. https://doi.org/10.1037/h0087468

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H. M., & Wittmann, W. W. (2007). Individual Differences in Components of Reaction Time Distributions and Their Relations to Working Memory and Intelligence. *Journal of Experimental Psychology: General*, *136*(3), 414–429. https://doi.org/10.1037/0096-3445.136.3.414

Schmitz, F., & Voss, A. (2012). Decomposing task-switching costs with the diffusion model. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 222–250. https://doi.org/10.1037/a0026003

Schubert, A.-L., Frischkorn, G., Hagemann, D., & Voss, A. (2016). Trait Characteristics of Diffusion Model Parameters. *Journal of Intelligence*, *4*(3), 7. https://doi.org/10.3390/jintelligence4030007

Schwarzkopp, T., Mayr, U., & Jost, K. (2016). Early selection versus late correction: Age-related differences in controlling working memory contents. *Psychology and Aging*, *31*(5), 430–441. https://doi.org/10.1037/pag0000103

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure?

Predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, *5*(JUL), 1–10. https://doi.org/10.3389/fpsyg.2014.00772

Sharma, D., Booth, R., Brown, R., & Huguet, P. (2010). Exploring the temporal dynamics of social facilitation in the Stroop task. *Psychonomic Bulletin & Review*, *17*(1), 52–58. https://doi.org/10.3758/PBR.17.1.52

Sherif, M. (1936). The psychology of social norms. Harper.

Sherman, J. W., Macrae, C. N., & Bodenhausen, G. V. (2000). Attention and Stereotyping: Cognitive Constraints on the Construction of Meaningful Social Impressions. *European Review of Social Psychology*, *11*(1), 145–175. https://doi.org/10.1080/14792772043000022

Sherman, J. W., Stroessner, S. J., Loftus, S. T., & Deguzman, G. (1997). Stereotype suppression and recognition memory for stereotypical and nonstereotypical information. *Social Cognition*, *15*(3), 205–215. https://doi.org/10.1521/soco.1997.15.3.205

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127–190. https://doi.org/10.1037/0033-295X.84.2.127

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor, & T. G. Reeve (Eds.), *Stimulus-response compatibility: An integrated perspective* (pp. 31−86). Amsterdam: North-Holland.

Sinclair, L., & Kunda, Z. (1999). Reactions to a Black professional: Motivated inhibition and activation of conflicting stereotypes. *Journal of Personality and Social Psychology*, *77*(5), 885–904. https://doi.org/10.1037/0022-3514.77.5.885

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of Measuring Recognition Memory: Applications to Dementia and Amnesia, *117*, 34–50.

Spence, K. W. (1956). *Behavior theory and conditioning*. Yale University.

Spence, K. W., Farber, I.E., McFann, H.H. (1956). The relation of anxiety (drive) level to performance in competitional and non-competitional paired-associates learning. *Journal of Experimental Psychology, 52*(5), 296–305. doi:10.1037/h0043507

Spencer, S. J., Fein, S., Wolfe, C., Fong, C., & Duinn, M. (1998). Stereotype activation under cognitive load: the moderating role of self-image threat. *Personality and Social Psychology Bulletin, 24*, 1139–1152. https://doi.org/10.1177/01461672982411001

Steinborn, M. B., & Huestegge, L. (2019). Socially alerted cognition evoked by a confederate's mere presence: analysis of reaction-time distributions and delta plots. *Psychological Research*. https://doi.org/10.1007/s00426-019-01143-z

Stewart, B. D., von Hippel, W., & Radvansky, G. A. (2009). Age, Race, and Implicit Prejudice. *Psychological Science*, *20*(2), 164–168. https://doi.org/10.1111/j.1467-9280.2009.02274.x

Strauss, B. (2002). Social facilitation in motor tasks: a review of research and theory. *Psychology of Sport and Exercise*, *3*(3), 237–256. https://doi.org/10.1016/S1469-0292(01)00019-X

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Strube, M. J. (2005). What Did Triplett Really Find? A Contemporary Analysis of the First Experiment in Social Psychology. *The American Journal of Psychology*, *118*(2), 271–286. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/15989124

Swets, J. A. (ed.) (1964). Signal detection and recognition by human observers. New York: Wiley

Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology*, *36*(7), 778–793. https://doi.org/10.1037/0022-3514.36.7.778

Thiessen, D. D. (1964). Population density, mouse genotype, and endocrine function in behavior. *Journal of Comparative and Physiological Psychology*, *57*(3), 412–416. https://doi.org/10.1037/h0039847

Thomas, S., Skitka, L., Christen, S., & Jurgena, M. (2002). Social facilitation and impression formation. *Basic and Applied Social Psychology, 24*(1), 67–70. https://doi.org/10.1207/S15324834BASP2401_6

Tolman, C. W. (1967). The Feeding Behaviour of Domestic Chicks as a Function of Rate of Pecking By a Surrogate Companion. *Behaviour*, *29*(1), 57–62. https://doi.org/10.1163/156853967X00262

Tomaka, J., Blascovich, J., Kibler, J., & Ernst, J. M. (1997). Cognitive and physiological antecedents of threat and challenge appraisal. *Journal of Personality and Social Psychology*, *73*(1), 63–72. https://doi.org/10.1037/0022-3514.73.1.63

Travis, L. E. (1925). The effect of a small audience upon eye-hand coordination. *Journal of Abnormal Psychology and Social Psychology*, *20*(2), 142–146. https://doi.org/10.1037/h0071311

Treisman, A. (1988). Features and Objects: The Fourteenth Bartlett Memorial Lecture. *The Quarterly Journal of Experimental Psychology Section A*, *40*(2), 201–237. https://doi.org/10.1080/02724988843000104

Triplett, N. (1898). The dynamogenic factors in pacemaking and competition. *The American Journal of Psychology*, *9*(4), 507–533. Retrieved from http://www.jstor.org/stable/10.2307/1412188

Uziel, L. (2007). Individual differences in the social facilitation effect: A review and meta-analysis. *Journal of Research in Personality, 41*(3), 579–601. https://doi.org/10.1016/j.jrp.2006.06.008

Uziel, L. (2016). Alone, Unhappy, and Demotivated: The Impact of an Alone Mind-Set on Neurotic Individuals Willpower. *Social Psychological and Personality Science*, *7*(8), 818-827. https://doi.org/10.1177/1948550616657597

Vaughan, G. and Hogg, M. A. (2005). *Introduction to social psychology.* 4th ed. Frenchs Forest, Sydney, Australia: Pearson Education Australia.

Voss, A., & Voss, J. (2007). Fast-dm: a free program for efficient diffusion model analysis. *Behavior Research Methods*, *39*(4), 767–775. https://doi.org/10.3758/BF03192967

Voss, A., Nagler, M., & Lerche, V. (2013). Diffusion models in experimental psychology: A practical introduction. *Experimental Psychology*, *60*, 385–402. https://doi.org/10.1027/1618-3169/a000218

Voss, A., Rothermund, K., Gast, A., & Wentura, D. (2013). Cognitive processes in associative and categorical priming: A diffusion model analysis. *Journal of Experimental Psychology: General*, *142*(2), 536–559. https://doi.org/10.1037/a0029459

Voss, Andreas, Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: an empirical validation. *Memory & Cognition*, *32*(7), 1206–1220. https://doi.org/10.3758/BF03196893

Voss, Andreas, Voss, J., & Lerche, V. (2015). Assessing cognitive processes with diffusion model analyses: A tutorial based on fast-dm-30. *Frontiers in Psychology*, *6*(MAR), 1–14. https://doi.org/10.3389/fpsyg.2015.00336

Wagenmakers, E. J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language*, *58*(1), 140–159. https://doi.org/10.1016/j.jml.2007.04.006

Wagstaff, G. F., Wheatcroft, J., Cole, J. C., Brunas-Wagstaff, J., Blackmore, V., & Pilkington, A. (2008). Some cognitive and neuropsychological aspects of social inhibition and facilitation. *European Journal of Cognitive Psychology*, *20*(4), 828–846. https://doi.org/10.1080/09541440701469749

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, *143*(5), 2020–2045. https://doi.org/10.1037/xge0000014

Wheeler, L., & Davis, H. (1967). Social disruption of performance on a DRL schedule. *Psychonomic Science*, *7*(7), 249–250. https://doi.org/10.3758/BF03331100

Whyte, G. (1993). Escalating Commitment in Individual and Group Decision Making: A Prospect Theory Approach. *Organizational Behavior and Human Decision Processes*, *54*(3), 430–455. https://doi.org/10.1006/obhd.1993.1018

Wühr, P., & Huestegge, L. (2010). The Impact of Social Presence on Voluntary and Involuntary Control of Spatial Attention. *Social Cognition*, *28*(2), 145–160. https://doi.org/10.1521/soco.2010.28.2.145

Young, A. W., Hellawell, D. J., & Hay, D. C. (1987). Configural information in face perception. *Perception 16*(6), 747-759. https://doi.org/ 10.1068/p160747

Yu, R. F., & Wu, X. (2015). Working alone or in the presence of others: exploring social facilitation in baggage X-ray security screening tasks. *Ergonomics*, *58*(6), 857–865. https://doi.org/10.1080/00140139.2014.993429

Zaffy, D. J., & Bruning, J. L. (1966). Drive and the range of cue utilization. *Journal of Experimental Psychology, 71*(3), 382–384. https://doi.org/10.1037/h0022963

Zajonc, R. B. (1965). Social facilitation. *Science*, *149*(3681), 269–274. Retrieved from http://www2.psych.ubc.ca/~schaller/Psyc591Readings/Zajonc1965.pdf

Zajonc, R. B. (1980). Compresence. In P. B. Paulus (ed.), *Psychology of Group Influence*. Hillside, NJ: Erlbaum.

Zajonc, R. B., Heingartner, A., & Herman, E. M. (1969). Social enhancement and impairment of performance in the cockroach. *Journal of Personality and Social Psychology*, *13*(2), 83–92. https://doi.org/10.1037/h0028063

Zajonc, R. B., & Sales, S. M. (1966). Social facilitation of dominant and subordinate responses. *Journal of Experimental Social Psychology*, *2*(2), 160–168. https://doi.org/10.1016/0022-1031(66)90077-1

Zentall, T. R., & Levine, J. M. (1972). Observational Learning and Social Facilitation in the Rat. *Science*, *178*(4066), 1220–1221. https://doi.org/10.1126/science.178.4066.1220

Attachments

Output 1.1.1- Experiment 1- Self-reports- Independent Sample T-Test

	statistic	df	p	Mean difference(PO-NPO)	Cohen's d
accompanied	-1.878	122	0.063	-0.5806	-0.3373
evaluated	0.871	122	0.385	0.3226	0.1565
observed	-1.306	122	0.194	-0.3226	-0.2346
demanding	-2.441	122	0.016	-0.5484	-0.4384
engagement	-0.340	122	0.735	-0.0645	-0.0610
motivation	0.516	122	0.606	0.1129	0.0928
attention	1.118	122	0.266	0.2258	0.2009
difficulty	0.976	122	0.331	0.2258	0.1753
positive-negative	0.160	122	0.873	0.0484	0.0287
sad-happy	-0.591	122	0.555	-0.1290	-0.1062
well-bad	0.603	122	0.548	0.1613	0.1083
bored-alerted	-0.166	122	0.868	-0.0484	-0.0298
rested-tired	-1.752	122	0.082	-0.4677	-0.3147
tense-relaxed	1.586	122	0.115	0.4355	0.2848

Output 1.1.2- Experiment 1 - Self-reports- Descriptive Statistics

	NPO mean(SD)	PO mean (SD)
accompanied	3.71(1.86)	4.29 (1.57)
evaluated	4.10(2.06)	3.77 (2.06)
observed	1.82 (1.20)	2.15 (1.53)
demanding	5.15 (1.35)	5.69 (1.14)
engagement	5.82 (1.01)	5.92(1.11)
motivation	5.34 (1.16)	5.23 (1.27)
attention	5.77 (0.95)	5.55 (1.28)
difficulty	3.29 (1.21)	3.06 (1.37)
positive-negative	3.66 (1.70)	3.61 (1.67)
sad-happy	4.66 (1.21)	4.79 (1.22)
well-bad	3.23 (1.40)	3.06 (1.58)
bored-alerted	4.50 (1.65	4.55 (1.60)

	NPO mean(SD)	PO mean (SD)
rested-tired	5.00 (1.64)	5.47 (1.31)
tense-relaxed	3.84 (1.52)	3.40 (1.54)

Output 1.2 - Experiment 1 -Error Rates - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,123	0.17	2.97	.09	.02
Target	1,123	0.00	0.04	.84	.00
Prime	1,123	0.02	7.90	>.01	.06
Target*Prime	1,123	0.09	22.59	>.001	.16
Target*Social Condition	1,123	0.01	0.79	.37	.01
Prime*Social Condition	1,123	0.00	0.43	.51	.00
Target*Prime*Social Condition	1,123	0.00	0.29	.59	.00

Output 1.3 - Experiment 1 -PDP C Component - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,123	0.35	2.97	.09	.02
Prime	1,123	0.03	7.90	>.01	.06
Prime*Social Condition	1,123	0.00	0.43	.51	.00

Output 1.4 - Experiment 1 -PDP A Component - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,123	0.01	3248.21	.58	.00
Prime	1,123	0.20	14.77	>.001	.11
Prime*Social Condition	1,123	0.01	0.54	.47	.00

Output 1.5 - Experiment 1 -Correct Reaction Times - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,123	63689	7.30	<.001	.06
Target	1,123	1678	5.81	.02	.05
Prime	1,123	216	1.42	.23	.01
Target*Prime	1,123	1026	8.17	>.01	.06
Target*Social Condition	1,123	123	0.42	.52	.00
Prime*Social Condition	1,123	115	0.76	.39	.01
Target*Prime*Social Condition	1,123	41	0.33	.58	.00

Output 1.6 - Experiment 1 -CDF - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,123	0.02	0.27	.60	.00
Bins	1,123	0.09	1.96	.10	.02
Social Condition* Bins	1,123	0.07	1.54	.19	.01

Output 2.1- Experiment 2- d600- Independent Sample T-Test

	statistic	df	p	Mean difference (NPO-PO)	Cohen's d
D600	-2.28	111	0.03	-0.14	.43

Output 2.2- Experiment 2- Error Rates- Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,111	0.00	66.30	.71	.00
Block	1,111	0.00	0.68	.41	.01
Block*Social Condition	1,111	0.00	0.75	.39	.01

Output 2.3 - Experiment 2 -PDP Components - Independent Sample T-Test

	statistic	df	p	Mean difference (NPO-PO)	Cohen's d
C Component	0.37	111	.71	.00	.01
A Component	-1.31	111	.20	-0.04	.20

Output 2.4- Experiment 2- Correct Reaction Times - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,111	295867	2.70	.10	.02
Block	1,111	647043	16.24	.<.001	.13
Block*Social Condition	1,111	115554	2.90	.09	.03

Output 2.5 - Experiment 2 -CDF - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,111	674066	3.31	.07	.03
Bins	1,111	494576	14.26	>.001	.11
Social Condition* Bins	1,111	106589	3.07	>.01	.03

Output 3.1.1 - Experiment 1 -Drift-Rates (Correct-Incorrect Thersholds) - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,123	5.99	2.00	.16	.02
Type of Trial	1,123	1.26	2.93	.09	.02
trial*Social Condition	1,123	0.00	2.90	.96	.00

Output 3.1.2 - Experiment 1 -Relative Starting-point (Correct-Incorrect Thersholds) - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,123	0.14	0.50	.48	.00
Type of Trial	1,123	0.01	1.14	.29	.01
trial*Social Condition	1,123	0.00	0.80	.79	.00

Output 3.1.3 - Experiment 1 -Threshold Separation (Correct-Incorrect Thersholds) - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,123	0.06	3.00	.09	.02
Type of Trial	1,123	0.00	0.36	.55	.00
trial*Social Condition	1,123	0.02	3.50	.06	.03

Output 3.1.4 - Experiment 1 -Nondecision Time (Correct-Incorrect Thersholds) - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,123	0.06	6.20	.01	.05
Type of Trial	1,123	0.00	0.49	.50	.00
trial*Social Condition	1,123	0.00	0.20	.66	.00

Output 4.1.1 - Experiment 1 -Drift-Rates (Gun-Tool Thresholds) - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,123	16.90	1.84	.18	.02
Target	1,123	11.85	18.45	>.001	.13
Prime	1,123	2.71	10.89	>.01	.08
Target*Prime	1,123	4.03	13.51	.02	.10
Target*Social Condition	1,123	0.06	0.08	.77	.00
Prime*Social Condition	1,123	0.02	0.08	.78	.00
Target*Prime*Social Condition	1,123	0.01	0.02	.89	.00

Output 4.1.2 - Experiment 1 -Relative Starting-point (Gun-Tool Thresholds) - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,123	0.00	0.12	.73	.00
Prime	1,123	0.00	0.17	.68	.00
Prime*Social Condition	1,123	0.00	0.11	.74	.00

Output 4.1.3 - Experiment 1 -Threshold Separation (Gun-Tool Thresholds) - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,123	0.06	5.87	.02	.05
Prime	1,123	0.00	0.00	.97	.00
Prime*Social Condition	1,123	2.13	2.131	.15	.02

Output 4.1.4 - Experiment 1 -Nondecision Time (Gun-Tool Thresholds) - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,123	0.12	6.47	.01	.05
Target	1,123	0.00	4.34	.04	.03
Prime	1,123	0.00	0.62	.43	.01
Target*Prime	1,123	0.00	0.16	.69	.00
Target*Social Condition	1,123	0.00	0.12	.73	.00
Prime*Social Condition	1,123	0.02	1.20	.28	.01
Target*Prime*Social Condition	1,123	0.00	0.30	59	.00

Output 5.1- Experiment 3- d600- Independent Sample T-Test

	statistic	df	p	Mean difference (NPO-PO)	Cohen's d
D600	-1.84	93	0.03	-0.23	.37

Output 5.2- Experiment 3- Error Rates- Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,93	0.02	2.06	.16	.02
Block	1,93	0.19	93.83	<.001	.50
Block*Social Condition	1,93	0.00	1.74	.19	.02

Output 5.2- Experiment 3- Correct Reaction Times - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,93	00	0.00	1.00	.00
Block	1,93	11	0.08	.78	.00
Block*Social Condition	1,93	142	1.01	.32	.01

Output 5.3 - Experiment 3 -PDP Components - Independent Sample T-Test

	statistic	df	p	Mean difference (NPO-PO)	Cohen's d
C Component	1.44	93	.16	0.04	.03
A Component	0.46	93	.65	-0.01	.01

Output 5.4 - Experiment 3 -CDF with RTs - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,93	1885.39	1.13	.29	.01
Bins	5,465	2296.91	16.83	>.001	.15
Social Condition* Bins	5,465	101.80	0.75	.59	.00

Output 5.5 - Experiment 3 -CDF with ERs - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,93	0.04	1.77	.19	.02
Bins	5,465	0.06	4.59	>.001	.05
Social Condition* Bins	5,465	0.01	0.59	.71	.01

Output 5.6 - Experiment 1 -Drift-Rates - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,93	0.11	0.08	.77	.02
Blocks	1,93	21.16	26.50	>.001	.22
Social Condition* Block	1,93	2.74	3.43	.07	.04

Output 5.7 - Experiment 1 -Relative Starting-Point - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,93	0.09	2.33	.13	.02
Blocks	1,93	0.15	9.12	>.01	.09
Social Condition* Block	1,93	2.56	2.56	.11	.03

Output 5.8 - Experiment 1 -Threshold Separation - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,93	0.01	0.28	.60	.00
Blocks	1,93	0.11	7.05	>.01	.07
Social Condition* Block	1,93	0.01	0.66	.41	.01

Output 5.9 - Experiment 1 -Nondecision Times - Mixed ANOVA

	df	Mean Square	F	p	η^2_p
Social Condition	1,93	1779	0.71	.40	.00
Blocks	1,93	2712	4.58	.04	.05
Social Condition* Block	1,93	307	0.52	.47	.01