DOI : 10.35662/unine-these-3046

Université de Neuchâtel

Faculty of Humanities
Cognitive Science Center
Communication and Cognition Institute

From learning about social categories to holding stereotypes

Investigating the acquisition of stereotypes in childhood and their effects on adults' information processing

Magali A. Mari

Thesis submitted in partial fulfilment of the requirements for the degree of

Docteure en Sciences Cognitives
Orientation : Psychologie

Members of the Jury:

Prof. Dr. Fabrice Clément (supervisor), Université de Neuchâtel, Switzerland
Prof. Dr. Gil Diesendruck, Bar-Ilan University, Israel
Prof. Dr. Yarrow Dunham, Yale University, USA
Prof. Dr. Markus Paulus, LMU Munich, Germany

Defended on the 20th of July 2023

unine
Université de Neuchâtel

Faculté des Lettres et Sciences Humaines
Espace Tilo-Frey 1
CH - 2000 Neuchâtel
Tél. +41 32 718 1700
secretariat.lettres@unine.ch

IMPRIMATUR

La Faculté des lettres et sciences humaines de l'Université de Neuchâtel, sur les rapports de M. Fabrice Clément, directeur de thèse, professeur, Université de Neuchâtel ; M. Markus Paulus, professeur, Ludwig-Maximilians-Universität, München ; M. Gil Diesendruck, professeur, Bar-Ilan University, Israel : M. Yarrow Dunham, professeur associé, Yale University, New Haven (USA) autorise l'impression de la thèse présentée par Mme Magali Mari en laissant à l'autrice la responsabilité des opinions énoncées.

Neuchâtel, le 20 juillet 2023

Le doyen
Louis de Saussure

UNIVERSITÉ DE NEUCHÂTEL

Rectorat
Fbg de l'Hôpital 41
2000 Neuchâtel
Tél. +41 32 718 10 20
messagerie.rectorat@unine.ch

Déclaration sur l'honneur*

Par la présente, j'affirme avoir pris connaissance des documents d'information et de prévention du plagiat émis par l'Université de Neuchâtel et m'être renseigné-e correctement sur les techniques de citation.

J'atteste par ailleurs que le travail rendu est le fruit de ma réflexion personnelle et a été rédigé de manière autonome.

Je certifie que toute formulation, idée, recherche, raisonnement, analyse ou autre création empruntée à un tiers est correctement et consciencieusement mentionnée comme telle, de manière claire et transparente, de sorte que la source en soit immédiatement reconnaissable, dans le respect des droits d'auteur et des techniques de citations.

Je suis conscient-e que le fait de ne pas citer une source ou de ne pas la citer clairement, correctement et complètement est constitutif de plagiat.

Je prends note que le plagiat est considéré comme une faute grave au sein de l'Université. J'ai pris connaissance des risques de sanctions administratives et disciplinaires encourues en cas de plagiat (pouvant aller jusqu'au renvoi de l'université).

Je suis informé-e qu'en cas de plagiat, le dossier sera automatiquement transmis au rectorat.

Au vu de ce qui précède, **je déclare sur l'honneur ne pas avoir eu recours au plagiat ou à toute autre forme de fraude.**

Nom : Mari

Prénom : Magali

Cursus : Doctorat

Faculté d'inscription : Faculté des Lettres et Sciences Humaines

Lieu et date : Delémont, le 20 août 2023

Signature : 

****

Ce formulaire doit être dûment rempli par tout étudiant ou toute étudiante rédigeant un travail substantiel (notamment un mémoire de bachelor ou de master) ou une thèse de doctorat. Il doit accompagner chaque travail remis au professeur ou à la professeure.

*Formulaire largement inspiré de la Directive de la direction 0.3 bis, intitulée Formulaire Code de déontologie en matière d'emprunts, de citations et d'exploitation de sources diverses, de l'Université de Lausanne, du 23 avril 2007 et adapté aux besoins de l'Université de Neuchâtel.

Summary

Assuming that boys are better at math than girls, expecting that Swiss people love to eat chocolate, or inferring that senior citizens are not at ease with new technologies, all reflect stereotypical expectations about specific social categories. The term "stereotype" refers to shared sets of expectancies about the likely personality, behaviors, preferences, or physical features of social category members. From a cognitive perspective, stereotypes are energy saving devices – mental shortcuts – that allow perceivers to spare processing costs when navigating their complex social environment. Yet, inferring people's preferences, behaviors, or personality from their category membership can also lead to incorrect predictions, to offensive assumptions, or even to discriminatory behaviors. The present dissertation aims to shed light on the mechanisms by which we form expectations about social categories, and how these expectations impact information processing.

Specifically, the first part of this dissertation provides a developmental perspective on how children learn about social categories. It specifically examines how children learn to make inferences about social categories, and the conditions under which children start attributing properties to social category members. The second part of this dissertation turns to the resulting stereotypes that adults hold about social categories, and how these stereotypes modulate information processing. This part focuses on the mental processes that underlie stereotyping and assesses how the processing of written texts is affected by items that confirm *versus* contradict stereotypical expectations. Together, this dissertation provides a cognitive and developmental perspective on the acquisition of stereotypes and their later effects on information processing. In doing so, this dissertation will hopefully bring a better comprehension of the foundational cognitive processes that underlie stereotyping and its consequences.

Résumé

S’attendre à ce que les garçons soient meilleurs en mathématiques que les filles, supposer que les Suisses aiment manger du chocolat ou déduire que les seniors ne sont pas à l’aise avec les nouvelles technologies, tout cela reflète des attentes stéréotypées à propos de certaines catégories sociales. Le terme stéréotype fait référence à un ensemble d’attentes partagées qui permet de prédire la personnalité, les comportements, les préférences ou les caractéristiques physiques des membres d’une catégorie sociale. D’un point de vue cognitif, les stéréotypes sont des raccourcis mentaux, qui permettent d’économiser de l’énergie et de limiter les coûts de traitement d’informations complexes issues de notre environnement social. Cependant, déduire les préférences, les comportements ou la personnalité d'une personne à partir de son appartenance à une catégorie sociale peut également conduire à des prédictions incorrectes, à des suppositions offensantes, voire à des comportements discriminatoires. L’objectif de cette recherche de doctorat est de mettre en lumière les mécanismes par lesquels nous formons des attentes à l’égard des catégories sociales et à la manière dont ces attentes influencent le traitement de l'information.

La première partie de la thèse offre une perspective développementale et expose la façon dont les enfants acquièrent des stéréotypes. Dans cette première partie, il sera question d’examiner comment les enfants apprennent à faire des inférences à propos de catégories sociales, et de déterminer les conditions dans lesquelles ces derniers commencent à attribuer des caractéristiques spécifiques à des membres de certaines catégories sociales. Dans la deuxième partie, l’accent sera mis sur les stéréotypes que les adultes peuvent avoir vis-à-vis de certains groupes sociaux et à la manière dont ces stéréotypes affectent le traitement de l'information. Cette partie se concentre sur les processus mentaux qui sous-tendent les stéréotypes et analyse comment le traitement d’informations écrites est affecté par des éléments qui confirment ou contredisent des attentes stéréotypées. Dans son ensemble la recherche présentée dans cette thèse de doctorat propose une perspective cognitive et développementale de l’acquisition des stéréotypes et de leurs effets sur le traitement de l'information, offrant ainsi une meilleure compréhension des processus cognitifs fondamentaux qui sous-tendent les stéréotypes et leurs conséquences.

Keywords / Mots-clés

English: social cognition; developmental psychology; stereotypes; social inferences; social categories; information processing.

Français : cognition sociale ; psychologie développementale ; stéréotypes ; inférences sociales ; catégories sociales : traitement de l'information.

Acknowledgements

So here I am, writing the acknowledgements for my PhD dissertation… a point I have been hoping to reach since I started my PhD six years ago. Six years of passion, discoveries, learning, friendships, and collaborations, but also so many mixed feelings: hope, fear, love, worry, doubt, fascination, but most of all courage and perseverance. I would not be at this stage of my career without the help of important people who are so dear to me.

First, I would like to sincerely thank my supervisor, Professor Fabrice Clément who opened me to the field of cognitive science. Eleven years ago, I started a Bachelor degree at the University of Neuchâtel, thinking I would leave academia right after graduation. Yet, here was this outstanding professor in communication sciences, whose lectures were so fascinating and captivating that I was upset for not being able to write down every single word he said. This increasing interest in Fabrice's field of research led me to enroll in the newly opening Master in Cognitive Science, and my fascination with his work continued to grow, to the point where I started dreaming of pursuing a PhD degree under his supervision. My dream came true when Fabrice offered me the position of PhD assistant within his team. Working under his supervision has been incredibly stimulating, opening me up to a vast field of research, bridging influences from developmental psychology, sociology, neuroscience, biology, and philosophy. His consistent enthusiasm had so much to do with getting my research going on, and our numerous conversations profoundly helped me to carry on with my doctorate research. I would never thank him enough for helping me in the realization of the studies presented in this dissertation, for having given me so many insightful feedbacks and constructive advices, and for inspiring me for the last ten years.

Over the years spent working on this PhD thesis, a wide network of colleagues has played a major role in the realization of the articles composing this dissertation and in the ultimate completion of this doctorate research. A particular thank to Misha-Laura Müller, with whom I worked in tight collaboration for the last four years. Her patience and dedication, as well as her clever and sharp ideas, contributed significantly to the writing of the second part of this dissertation. Professor Markus Paulus and his team at LMU Munich have significantly advanced my knowledge in developmental psychology, and have provided me with a great deal of advice and guidance in all the reflections that accompanied the developmental part of the dissertation. Without their help and valuable insights, I would never have been able to

successfully to carry out a study with children during a pandemic. For their support and wonderful presence during my PhD, my gratitude goes to my colleagues, Adrián Petschen, Léa Farine, Fabienne Voirin, Cristina Grisot, Mélinda Pozzi, Diane Liberatore, Noémie Treichel, and Diana Mazzarella. I also would like to sincerely thank Professor Yarrow Dunham and Professor Gil Diesendruck for the time they dedicated to providing me with clear and helpful comments and feedbacks that helped me improve previous versions of this dissertation. It is an honor for me to have these eminent researchers in my PhD committee.

All PhD thesis involves many hours of thinking and writing that would have felt very lonely and insurmountable without my family and friends to cheer me up and remind me to enjoy life. My eternal gratitude goes to the Green House Clan, whose renown members have supported me since childhood, namely my sister Stéphanie Mari, my mother Anita Comment, and my stepfather Gérard Comment, to my family across the mountain, my father Jean-René Mari, my stepmother Stéphanie Mari-Cuenat, and my little sisters Théa Mari and Chiara Mari, to my friends, in particular Alexandre Rossé and Vincent Senn, and, of course, to the inhabitants of the Principality of Péca, Valérie Guéniat and Cédric Guéniat, my heart-adopted parents, who have been so compassionate and have always kept faith in me.

I saved the best for last: There are no words to do justice to the thanks my partner Joé Gueniat deserves. He has seen me through every state, living all the joys and good news but also the suffering that have accompanied this doctorate. I thank him infinitely for his patience and unfailing support, for reassuring and reasoning with me when I was in deep stress. His precious advice has guided me in every single step of this PhD degree, so much that I would never have had the force to get through it without him. I am incredibly lucky to share my life with such a wonderful and brilliant man, and I cannot wait to see what awaits us in the future.

Index

1. General introduction

The way people readily make inferences about others has always intrigued me. I remember once going to the florist to find a present for a friend who just moved into a new house. Explaining to the florist that I was looking for a plant to bring to the housewarming, she went to the back of her boutique and came back with a plant, saying enthusiastically "I am sure she will love this fine and delicate pink dracaena." I could not help but smile at her assumptions. For one thing my friend was a man, but worst of all, why would "she" love pink fine and delicate plants? Everyone can remember situations like this one. And while they make for funny anecdotes, they also bring to light a general tendency to rapidly generate expectations – probably incorrect ones – about others from scarce information. This dissertation examines how we come to form expectations about others in childhood and how these expectations shape the way we process information as adults.

Social interactions are pervasive in human life. From our first day of life, we are immersed into a rich and complex social environment. Classifying people into categories is an efficient way of structuring and simplifying the amount of information we encounter in this rich environment (Bigler & Liben, 2006; Rhodes, 2013; S. J. Sherman et al., 2013; Tajfel, 1969). Based on social category membership, we can predict and explain new individuals' actions by drawing on past experiences (Baron, 2009; Rhodes, 2013; Rhodes & Baron, 2019). The florist might have come to infer that women tended to choose pink and delicate flowers from her client experience. Yet, classifying people into categories and inferring their preferences, behaviors, or personality from their category membership can also lead to incorrect predictions, to offensive assumptions, or even to discriminatory behaviors. Indeed, these expectations based on people's category membership have been documented to have profound influences in daily life, affecting friendship choices, hiring decisions, or even jury verdicts (Dovidio et al., 2002; Fiske & Taylor, 2013; Sidanius & Pratto, 1999). As such, social categorization has cognitive advantages, reducing the complexity of the social world and allowing us to quickly encode and retrieve information about people (Rhodes, 2013), but it also sets the ground for the development of stereotypes, which can in turn have serious consequences and lead to many social issues.

Making inferences about the likely properties of others based on their category membership is a first step toward holding stereotypical expectations. Stereotypes are defined

as mental constructs that reflect acquired knowledge or experiences about social categories, and that generate expectancies about the likely behaviors, preferences, psychological traits, or physical features of category members (D. Carlston, 2010; Dovidio, 1999; Over & McCall, 2018; Quadflieg & Macrae, 2011). From this perspective, stereotypes are forms of information that have been acquired over many years from language and social experiences that "cumulatively construct an overlearned repertoire of cultural expertise" about social categories (Greenwald & Banaji, 2017, p. 866). As argued by Clément (2010), stereotypes result from a kind of progressive "imprint" during the socialization process, whereby individuals assimilate the representations of their own group (see also McGarty et al., 2002a). The family and other socialization agents, as well as the media, will play a significant role on the contents of stereotypes, as some associations will appear more frequently than others depending on the cultural context (Bar-Tal, 1996; Jost & Hamilton, 2004). Think for example, of all the TV commercials that portrayed little girls playing with dolls and boys playing with trucks, of movies where the bad guys were Latinos and African Americans, or of your neighbor complaining about the increasing number of immigrants who recently moved in your town… This unsought cultural expertise is acquired in early childhood and "fills in" stereotypical representations that link social categories with certain attributes (Charlesworth et al., 2021; Greenwald & Banaji, 2017). It results that, upon encountering or hearing about social category members, a set of possibly erroneous expectations about the likely attributes of this social category appears automatically in our minds.

The present dissertation aims to shed light on the origin and manifestation of stereotypes. It offers a better understanding of the mechanisms by which we form expectations about social categories, and how these expectations impact information processing. The dissertation is articulated in two parts. The first part provides a developmental perspective on how children learn about social categories. It specifically examines the following questions: (1) How do children learn to make inferences about social categories?; and (2) Under which conditions do children start generalizing properties across members of the same social category? The second part turns to the resulting stereotypes that adults hold about social categories and how these stereotypes affect information processing. Specifically, this part addresses the following questions: (3) How and why does stereotyping stem, but yet differ, from social categorization processes?; and (4) How does information that confirms *versus* contradicts stereotypical expectations modulate the processing of written texts? Together, this

dissertation provides a cognitive and developmental perspective on the acquisition of stereotypes and their later effects on information processing.

1.1. Common processes involved in social categorization and stereotyping

A long tradition of research in social psychology has conceived stereotypes as closely linked to categorization process (Allport, 1954; D. Hamilton, 1981; Lepore & Brown, 1997; Tajfel, 1969). This assumption rests on the fact that stereotypes serve similar functions as general categorical knowledge, namely structuring and simplifying the important number of stimuli that we encounter (S. J. Sherman et al., 2013). By classifying people into a same category, one can predict the specific features of any one category member (Bruner, 1957; Cantor & Mischel, 1979). And this is exactly what stereotypes do: They generate shared sets of expectancies about the likely traits, roles, behaviors, preferences, or physical features of category members (Dovidio, 1999; Over & McCall, 2018; Quadflieg & Macrae, 2011). As such, the processes involved in social categorization are assumed to be largely related to those that underly stereotypical knowledge.

Classical accounts considered categories, both social and nonsocial, as being specified by a set of necessary and sufficient features common to all members of a given category (Bourne, 1966; Bruner et al., 1956), e.g., category *X* has features *a,* and *b*; if *y* does not have features *a* and *b*, then it is not a member of *X*. However, influential research, notably conducted by Rosch and colleagues, challenged this classical approach, showing that categories used in everyday experience often fail to fall under the rule of the "all-or-none" criterion of category membership (Rosch, 1978; Rosch et al., 1976; Rosch & Mervis, 1975). Rather, the seminal work of Rosch and colleagues offered a view of categories as specified by probabilistic distributions of features. Categories are organized around prototypical stimuli with category members varying from the central prototype, which corresponds to an abstract "average," a generalized representation, of the members of a category (Rosch & Mervis, 1975). According to this perspective, category members have *family resemblance* structures, such that category members have properties in common, but no set of properties is commonly held by *all* members (Rosch, 1978). Other perspectives emphasized the role of exemplars, defending the idea that categorization is based on the comparison of a stimulus with instances already stored in memory; if the stimulus fits the retrieved exemplars, it is ascribed to the category (Brooks, 1978; Medin & Schaffer, 1978; E. R. Smith & Zárate, 1992). Both accounts share the driving

principle that categorization is based on the degree of similarity between an item and other category members, either according to its resemblance with a prototype or with an exemplar (Bar-Tal, 1996; Jost & Hamilton, 2004; Murphy, 2016).

Social categories follow the same principles, namely social category members are related by family resemblance rather than by fulfilling necessary and sufficient criteria. Within a given social category, some members will be considered as more typical than others. Typicality, in this sense, refers to the extent to which a person resembles a prototype – i.e., a summarized representation of characteristic features – or exemplars – i.e., actual people considered as the most representative of the category[1] (Bar-Tal, 1996; Fiske & Taylor, 2013; Rosch, 1978; E. E. Smith & Medin, 1981). For instance, one could ascribe a person to the social category "Swiss" from their resemblance with a prototype (e.g., punctual, good at skying, eat chocolate) or exemplars (e.g., Joël Dicker, Sarah Marquis). Once an individual has been ascribed to a social category, one can readily infer the likely properties of this individual by drawing on their past experience of the social category, i.e., known exemplars or related prototypical features. And this tendency is present very early; from preschool years, children are able to generalize the properties they learned about a social category to other individuals belonging to this social category (cf. Chapter 1 and Chapter 2). Stereotypes find their roots in these categorization processes: As one learns to associate some properties with a social category, one also starts to construct a set of stereotypical expectancies about the members of this social category (Dovidio, 1999).

1.2. Stereotypes as biased representations of social categories

Stereotypes are more than sets of descriptive features associated with category members. Rather, they reflect extreme representations of social categories that are rigid and resistant to change, that ignore within-category variations, and that can have serious social consequences (Foster-Hanson & Rhodes, 2022; Hilton & von Hippel, 1996). Some theoretical accounts proposed that stereotypes mainly emerge from limitations of the human capacity for processing information (Simon, 1955; Stroebe & Insko, 1989; Tajfel, 1969). When social agents are classified into categories, perceivers start to view members of the same social category as

[1] Note that while the exemplar and prototype approaches have often been contrasted, it is now well established that people rely on both exemplars and prototypes to categorize people (e.g., Cohen & Basu, 1987; Fiske & Taylor, 2013; Johansen & Kruschke, 2005; Medin et al., 1984; E. E. Smith & Medin, 1981; Yamauchi & Markman, 1998).

more similar, and members of different social categories as more different from each other (Eiser & Stroebe, 1972; Ford & Stangor, 1992a; Tajfel & Wilkes, 1963). These within- and between-group distortions tend to generalize to additional features, beyond those that differentiated social categories in the first place (Allport, 1954; Gaertner & Dovidio, 2004), and these features are ascribed indiscriminately to *all* members of a social category (Jost & Hamilton, 2004). As such, stereotypes have been considered as oversimplified representations of social categories that ignore the complexity and diversity of people.

Other theoretical accounts emphasized the role of affective and motivational factors in the emergence of stereotypes (Allport, 1954; Greenwald et al., 2002; Tajfel & Turner, 1979; Walton & Banaji, 2004). Because stereotypes refer to categories of people, they are self-relevant and socially relevant in ways that general categorical knowledge is not (S. J. Sherman et al., 2013). While categorizing others, we always also implicitly categorize ourselves, and come to divide the world into ingroups, the people to whom we belong, and outgroups, those to whom we do not belong (P. M. Brown & Turner, 2002; Dunham, 2018). As one highly identifies with a group, the ingroup's welfare and image become closely related to one's own well-being and self-esteem (Cvencek et al., 2021; Greenwald et al., 2002; Tajfel & Turner, 1979). This, in turn, motivates people to view their ingroup positively, which might be achieved by attributing less favorable properties to outgroup members compared to ingroup members (Dovidio, 1999; Judd & Park, 2004; Stroebe & Insko, 1989). If the properties associated with a social category are perceived as negatively valenced, then stereotypes may take on an affective aspect and become source of negative attitudes – or prejudices – toward the social category (Allport, 1954). Stereotypes, in this sense, are biased representations of social categories that can operate in the service of intergroup bias and prejudices. Chapter 3 provides a detailed review of the above cognitive bias and social motives that underly stereotypes.

1.3. Underlying processes of stereotypes

Early research on stereotypes mainly used explicit self-reports, such as asking participants to rate category members on descriptive adjective scales (e.g., Greenwald et al., 1998) or to decide which candidate they would hire for a stereotype-related job (e.g., Glick et al., 1988). These measures relied on participants' capacity for introspective access to their stereotypes, but people do not always know what is on their mind (Kurdi & Banaji, 2022). Starting from the 1980s, a new era of research in social psychology started to consider indirect

methods to assess stereotypes, relying on response times, accuracy rates or other responses that are likely to bypass deliberate control (e.g., Fazio et al., 1995; Greenwald et al., 1998; Nosek & Banaji, 2001). These methods were promising as they solve the problem that people might misreport their stereotypes because of social desirability (Blair, 2002) or because they might even be unaware of them (Nosek et al., 2011). A flourishing number of research soon followed, providing convincing evidence that stereotyping processes could operate in unconscious and unintentional ways, as well as consciously. This dual view constituted an important conceptual root for the understanding of the processes involved in stereotyping.

Building on classical dual-process conceptions of mental operations (e.g., Anderson & Bower, 1973; Collins & Loftus, 1975; Wyer & Carlston, 2018), stereotypes have been conceived as associated networks of concepts stored in semantic memory in a dormant state until they are activated in a given situation (Devine, 1989; Gilbert & Hixon, 1991; Macrae & Bodenhausen, 2000). Upon encountering or referring to a social category (e.g., woman), a set of related, possibly erroneous, concepts (e.g., caring, delicate, mother) activates automatically in people's minds (Gilbert & Hixon, 1991; Krieglmeyer & Sherman, 2012; Quadflieg & Macrae, 2011). The activation of stereotypes is thus considered as automatic because it happens relatively quickly and outside of awareness. Under some conditions, such as ego-protective motivations (Fein & Spencer, 1997), personal tendency to hold essentialist beliefs (Levy et al., 1998), or contextual and cognitive demands (van Knippenberg et al., 1999), people might use the activated stereotypical expectations to judge or evaluate others (Macrae & Bodenhausen, 2000; McGarty, 2002; Quadflieg & Macrae, 2011). The term *implicit stereotypes* has been coined to refer to the first process, i.e., the automatic activation of associations about social categories, whereas the term *explicit stereotypes* reflects the application of stereotypes, i.e., consciously endorsed beliefs about social categories.

At the end of the 20th century, research in social psychology started to show that stereotypes were more often subtle, automatic, and unintentional, than controlled and blatantly expressed (Fiske & Taylor, 2013; Kurdi & Banaji, 2022). While implicit forms of stereotypes were frequently documented, openly negative attitudes and explicit stereotypes started to dissipate (Crosby et al., 1980; Saucier et al., 2005). For this reason, measures of implicit stereotypes gained increased interest because they offered a potential explanation for why stereotypes and prejudices persisted, despite growing effort and attention to eliminate them (Charlesworth & Banaji, 2019). The studies presented in Chapter 4 and Chapter 5 build on this

tradition, insofar as they probe the automatic activation of stereotypes by using indirect measures.

1.4. Methodological considerations

1.4.1. Toward more openness and transparency

In the early 2010s, hundreds of researchers around the world started to conduct replications of classical and contemporary findings in social psychology. Between 2015 and 2018, three influential articles were published, revealing that more than half of the replicated studies failed to reproduce the original results (Camerer et al., 2018; R. A. Klein et al., 2018; Open Science Collaboration, 2015). These disturbing findings led to the beginning of a "replication crisis" in social psychology and raised important concerns about the methodological approach employed in the past. It became obvious that the field was facing pervasive problems regarding how data were reported, analyzed, and possibly selected for publication (Renkewitz & Heene, 2019). The main issue came from problematic data analysis practices, such as formulating hypotheses from the analyzed data, selectively dropping or adding participants to reach significant results, collecting data from too small samples (Davis-Kean & Ellis, 2019) or samples exclusively composed of people from Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies (Arnett, 2008; Henrich et al., 2010; Thalmayer et al., 2021). Important initiatives, promoting the now well-known open science movement, were undertaken to improve the way studies are conducted, from the study design, through data analyses and results reporting, to the publication process. Inscribed within this evolving research context, the present dissertation has been strongly impacted by these upcoming changes.

From its beginning in 2017, my journey as a PhD student has been greatly influenced by open science practices. Throughout my doctorate research, I gradually integrated the recommendations to improve the studies' reproducibility and replicability. The first step was to make the collected data, study materials, and statistical analyses publicly available, notably at the Open Science Framework. In a second step, the studies' hypotheses and experimental design, as well as sampling and analysis plans, were pre-registered before data collection. The replication crisis in social psychology also had a major impact on the methodologies selected for the present doctorate research. Concerns about the replicability of priming studies led to a

certain reluctance to employ those methods to study the automatic activation of stereotypes (Cesario, 2014; Kidder et al., 2018; Loersch & Keith Payne, 2016). The ongoing debate and controversies on implicit measures, such as the Implicit Association Test (Greenwald et al., 1998), the Evaluative Priming Task (Fazio et al., 1995), or the Go/No-go Association Task (Nosek & Banaji, 2001), also put into question the viability of these methods (Gawronski et al., 2020; Greenwald & Lai, 2020; Payne & Gawronski, 2010). Finally, the increasing need to replicate findings cross-culturally motivated the selection of protocols that were not bound to a specific cultural setting. Together, these factors played a decisive role in the methodologies adopted throughout the present research.

1.4.2. Selected methodologies

In the last 25 years, a wide area of research employed indirect measures to assess the automatic activation of stereotypes. As presented above, this line of research built on the limitations of explicit self-report measures, insofar as people might misreport their stereotypes because of social desirability (Blair, 2002), or because they underestimate or are unaware of them (Nosek et al., 2011). The basic idea behind indirect measures is to compare reaction times and/or error rates that occur when people quickly sort stimuli in different categories (e.g., classifying the stimulus "firefighter" in the category "women" or "men") (Kidder et al., 2018). A priming procedure might be employed, as in the Semantic Priming Task (Banaji & Hardin, 1996a; Wittenbrink et al., 1997), where participants need to classify a target (e.g., firefighter) after being presented with a prime that either matched stereotypical expectations (e.g., Joe) or not (e.g., Alice). These indirect measures are used to assess automatic stereotypical expectations by measuring, with reaction times, the strength of activation of associated concepts (e.g., firefighter and men) (Greenwald et al., 2003). These measures have been argued to capture representations that people might not be aware of and that operate outside of consciousness (Gawronski et al., 2020; Greenwald & Banaji, 1995). However, following the replication crisis in social psychology, an increasing number of issues and concerns were raised toward those methods and their broader theoretical and empirical implications. This growing number of critiques motivated the search for alternatives to measure the activation of stereotypical expectations.

The studies presented in Chapter 4 and Chapter 5 stemmed from an effort to find alternative methods to capture the automatic activation of stereotypes by drawing on research in psycholinguistics. Studies in this field have long established that reading a text involved a

wide range of cognitive processes oriented toward understanding the presented content, which can occur in rather discrete ways (Gygax et al., 2021; Perfetti & Stafura, 2014). While reading, one must not only visually process the written words but also understand their underlying meaning. To comprehend a text, readers combine elements that are derived explicitly from the text, as well as elements that are implicit, coming from their previously acquired knowledge (Elbro & Buch-Iversen, 2013; Graesser et al., 1994; Gygax et al., 2021; Kendeou et al., 2016). Importantly, upon reading a word or a sentence, related concepts are automatically activated in semantic memory (Gerrig & McKoon, 1998; O'Brien et al., 1998; Rapp & van den Broek, 2005; Rubio-Fernández, 2013). For instance, upon reading about "the firefighter," stereotypical representations of the protagonist will be activated from readers' memory, leading readers to hold expectations about the likely personality, behaviors, or physical features of "the firefighter" (Beukeboom & Burgers, 2019; O. Klein & Bernard, 2015). If readers encounter information that contradicts their expectations, they will need more time to integrate this unexpected item, resulting in an interruption of their normal reading pace (Gilbert & Hixon, 1991; Quadflieg & Macrae, 2011; Rees et al., 2020). Reading tasks could thus constitute a potential candidate to investigate how stereotypes automatically generate expectations about social agents. Deprived from sorting and priming procedures, reading tasks can also prove to be an interesting asset to the study of stereotype activation because they are close to natural settings and do not interrupt the ordinary flow of cognition (Coronel & Federmeier, 2016). The objective of Chapter 4 and Chapter 5 is to provide a detailed assessment of whether and how stereotypical expectations affect information processing by using reading tasks.

On a final note, the present doctorate research also adopted a developmental perspective to identify early-emerging, possibly foundational, processes that underlie the acquisition of stereotypes about social categories. Deciphering the age periods in, and conditions under, which children develop stereotypes can offer valuable insights into adults' stereotyping. These questions are addressed in Chapter 1 and Chapter 2. In addition, the study presented in Chapter 2 employed a protocol that was designed to enable the reproducibility and replicability of the study. Following the recommendations for more transparency and openness in research, the study design, sampling, and analysis plans were pre-registered before data collection. Moreover, the materials, stimuli, collected data and analyses' script have been made publicly available at the Open Science Framework. The manuscript has been written with RMarkdown (RStudio Team) to allow its full reproducibility and is freely available online. The study presented in Chapter 2, thus follows the most recent recommendations for transparency and the

latest open science practices, thereby aligning with political and ethical efforts to enhance the quality of research and its dissemination to all levels of the scientific community and the general public.

1.5. The following chapters

To address how children learn to make inferences about social categories, Chapter 1 reviews recent works in developmental psychology that investigated how children, between two and twelve years old, identify and draw inferences about novel social categories. This chapter compares the different kinds of inferences children draw about people based on their category membership, namely expectations about similarity, essentialist beliefs, deontic inferences, as well as affective responses to social category members. Chapter 1 details how children come to make each of the above inferences, delineates the needs for future research, and introduces key concepts and methodologies on which Chapter 2 relies.

Chapter 2 presents an experimental study assessing the conditions under which children start generalizing properties across members of the same novel social category. This study investigates two theoretical perspectives proposing that children may learn to associate properties with social categories from simple category labels and/or from their observation of regular behaviors. Chapter 2 also discusses how attributing properties to social category members sets the ground for the acquisition of stereotypes. Altogether, Chapter 2 provides key insights on the critical age periods in which children start to readily generalize properties to social categories, thereby contributing to the understanding of the central roles of language and observation in the development of stereotypes.

From research on how children learn to associate properties with social categories, Chapter 3 turns to the resulting stereotypes that adults hold about social categories. Specifically, this chapter first addresses the key question of how and why stereotyping differs from general categorical thinking by presenting cognitive bias and social motives that leads to the formation of biased, oversimplified, representations of social categories. Chapter 3 then presents how stereotypes might work as mental shortcuts that generate automatically expectancies about the likely properties of social categories. In doing so, Chapter 3 introduces classical methods that have been used to probe the automatic activation of stereotypes and discusses theoretical and methodological considerations relevant to the studies presented in the following chapters.

Chapter 4 reports an experimental study examining how stereotypical expectations about gender and nationalities affect information processing during reading. Across two experiments, this study shows that readers experience processing difficulties upon encountering stereotype-incongruent information, and that these difficulties are exacerbated by specific linguistic markers. By documenting that information violating a stereotype is immediately difficult to process, this study provides evidence for the automatic activation of stereotypical expectations about social categories during reading.

To further assess the automatic activation of stereotypical expectations during reading, Chapter 5 replicates the study presented in Chapter 4. Rather than using a self-paced reading task, the study employs an eye-tracking task, thereby offering a finer moment-by-moment analysis of information processing in a natural reading setting. In doing so, the study integrates theoretical perspectives on stereotyping and eye movement control, while specifying how and when the automatic activation of stereotypes modulates information processing during reading. Chapter 5 also discusses how the eye movement patterns observed in this study relate to the findings of research on stereotype processing using event-related brain potentials. Overall, Chapter 5 further details how stereotypical expectations modulate the processing of written information.

Together, this dissertation advances our understanding of how we form stereotypical expectations about others and how they shape the way we process information. I will draw on the research presented across the five chapters in this dissertation to specify how children learn to make inferences about social categories and how these inferences affect adults' information processing. Specifically, by examining how children learn to attribute properties to social categories, the first part of this dissertation aims to decipher the mediating factors of stereotype acquisition. The second part documents how information processing is modulated by stereotypical expectations, thereby specifying the nature of the cognitive processes involved in stereotyping. In doing so, this dissertation will hopefully bring a better comprehension of the foundational cognitive processes that underlie stereotyping and its consequences. And perhaps in doing so, this dissertation might contribute to inform future intervention programs designed to shape or mitigate the possible negative consequences of stereotypical expectations.

2. Chapter 1

How cues to social categorization impact children's inferences about social categories[2]

Abstract

Social categorization involves two crucial processes: First, children seek properties on which they can categorize individuals, i.e., they learn to form social categories; then children make inferences based on social category membership and might develop affective responses toward social categories. Over the last decade, a growing number of research in developmental psychology started to use novel social categories to investigate how children learn and reason about social categories. To date, three types of cues have been put forward as means to form social categories, namely linguistic, visual, and behavioral cues. Based on social category membership, children draw inferences about the shared properties of social category members and about how social category members ought to behave and interact with each other. With additional input, children might apply essentialist beliefs to social categories and develop affective responses toward social categories. This article aims to provide key insights on the development of stereotypes and intergroup biases by reviewing recent works that investigated how children learn to form novel social categories and the kind of inferences they make about these novel social categories.

2.1. Introduction

Social interactions are pervasive in our everyday life. We encounter countless people every day, by going to work, having a coffee, or taking a course at the gym. We are immersed in this rich and complex social environment from our first day of life. Classifying people into categories is an efficient way of structuring and simplifying the amount of information we encounter in the social world (Bigler & Liben, 2006; Rhodes, 2013; S. J. Sherman et al., 2013; Tajfel, 1969). In fact, social categorization allows children, like adults, to predict and explain

[2] Reprint of: Mari, M. A. (2022). How cues to social categorization impact children's inferences about social categories. *Acta Psychologica*, *229*(103707), 1–12. https://doi.org/10.1016/j.actpsy.2022.103707
This research was funded by the Swiss National Science Foundation, under the grant number P1NEP1_191722.

new individuals' actions by drawing on past experiences (Baron, 2009; Rhodes, 2013; Rhodes & Baron, 2019). Yet, social categorization also sets the ground for the development of stereotypes and intergroup biases (Bigler & Liben, 2006; Rhodes & Baron, 2019; Tajfel, 1970). As such, a precise understanding of the processes involved in social categorization can shed light on the origin of stereotypes and intergroup biases. Specifically, two processes are involved in social categorization: First, children seek properties on which they can categorize individuals, i.e., they learn to form social categories; then children make inferences based on social category membership and might develop affective responses toward social categories (Baron et al., 2014; Over & McCall, 2018; Shutts & Kalish, 2021). The present article aims to provide key insights on the development of stereotypes and intergroup biases by reviewing recent works that investigated how children learn to form novel social categories and the kind of inferences they make about these novel social categories.

From an evolutionary perspective, social categorization serves the adaptative purpose of detecting potential social partners, alliances, and coalitions (Kinzler et al., 2010; Spelke & Kinzler, 2007). Humans have thus developed a system for thinking about others as social category members, which is present early in ontogeny: Between 3 and 9 months, infants show looking preferences and categorization abilities (Charlesworth & Banaji, in press; Liberman et al., 2017). Based on social category membership, children are able to predict the properties and interactions of people belonging to the same social category (Baron, 2009; Liberman et al., 2017; Rhodes, 2013). Specifically, recent works in developmental psychology showed that children expect members of the same social category to share traits, activities, behaviors, physical attributes, and preferences (Over & McCall, 2018), that the properties marking social category members will endure over time (Liberman et al., 2017), and that social category members are obligated to one another, namely by cooperating and engaging in prosocial behaviors (Kaufmann & Clément, 2014; Rhodes & Chalik, 2013). Therefore, social categorization has cognitive advantages, reducing the complexity of the social world and allowing us to quickly encode and retrieve information about people (Rhodes, 2013), but it has also serious social and psychological drawbacks, namely intergroup biases.

The types of inferences children make about social categories are directly related to stereotypes and intergroup biases. Indeed, research in social psychology proposes that stereotypes correspond to the set of shared beliefs about the properties of social categories (O. Klein & Bernard, 2015), whereas attitudes are defined as positive or negative affective responses toward a social category (Kurdi, Mann, et al., 2019). Although stereotypes consist in

shared beliefs about a social category, they do not necessarily imply an affective aspect (e.g., one can hold a stereotype that gay couples have good taste in fashion *versus* believing that gay couples are deviant). Attitudes, however, involve a positive or negative evaluation (associating a category with good/bad), which can in turn lead to prejudices or even discrimination (e.g., favoring one category over another) (Fiske & Taylor, 2013). As previously mentioned, social categorization can lead to associating properties to social category members (which corresponds to forming stereotypes), and to developing a preference for some categories over others (which corresponds to developing positive and negative attitudes). Studying social categorization among children can reveal important aspects of how stereotypes, attitudes, and prejudices develop. As Devine (1989) argued, beliefs formed in childhood might be more rooted and thus less accessible to consciousness than recently acquired beliefs. Consequently, it is crucial to have a better understanding of how children infer social category's characteristics and develop affective responses toward a social category.

The present paper is articulated in two main sections, corresponding to the two processes involved in social categorization. The first part of the article presents the kinds of cues on which children rely to identify social categories. Previous theoretical accounts proposed that different cues can trigger social categorization, namely linguistic, visual, and behavioral cues (Bigler & Liben, 2006; Rhodes & Baron, 2019; S. J. Sherman et al., 2013). The second part of this article focuses on the types of inferences children draw based on social category membership. Research in developmental psychology showed that inferences can be about the social category's properties (Over & McCall, 2018), patterns of interpersonal interactions (Liberman et al., 2017), and essentialist beliefs about the social category (Rhodes, 2013). In addition, social categorization can also yield affective responses toward a social category (Spelke & Kinzler, 2007). This review focuses on studies using novel and abstract social categories as opposed to real-world social categories. This choice is motivated by the growing number of research using novel social categories over the last decade (with a few exceptions beforehand, for e.g., Ford & Stangor, 1992; Gregg et al., 2006; Levy et al., 1998). Moreover, investigating the processes of social categorization with novel social categories has the advantage of identifying more general aspects of social categorization that are not tied to cultural settings. As argued by Baron et al. (2014) using novel social categories "allows us to examine the abstract principles by which children create [and reason about] social categories in the first place." (p. 241) Because this review aims to propose opportunities for future research, it seems particularly interesting to focus on research with novel social categories as they uncover general processes and inferences

that might be tested and replicated in other cultures. The studies reviewed in the present article were thus selected on the basis of the following criteria: (a) experimental study, (b) use of novel social categories; (c) tested population included infants, preschoolers and/or school-aged children; (d) assessment of how social categories are formed (i.e., social category formation) and/or (e) assessment of social category inferences (i.e., similarity inferences, deontic inferences, essentialist beliefs, or affective responses), and (f) publication date comprised between January 2010 and December 2021. Table 1 provides an overview of the selected studies for the present review. Altogether, the present article delineates how specific cues to identify novel social categories can lead to different types of inferences, draws key conclusions about the conditions under which children learn and make inferences about novel social categories, and proposes opportunities for future research.

Table 1. Summary of research using novel social categories

Cues primarily tested in the study	Additional cues used in the experimental procedure	Process of social categorization under investigation	References
Linguistic (labels, generics, specifics)	Visual (clothes colors)	Social category formation	Rhodes, Leslie, Bianchi, et al. (2018)
Behavioral	Visual (clothes & hair colors)	Social category formation Similarity inference (activity)	Riggs (2019)
Visual (clothes, hat, skin colors) *versus* Linguistic (labels)		Social category formation Similarity inference (activity) Affective responses	Baron et al. (2014) Diesendruck & Weiss (2015)
Linguistic (labels, generics)		Similarity inference (physical features)	Diesendruck & Eldror (2011)
Linguistic (labels, specifics)	Visual (clothes style & colors)	Similarity inference (preferences, physical features, knowledge)	Kalish (2012); Soley (2019); Riggs et al. (2014)
Linguistic (labels, generics)	Visual (clothes colors)	Similarity inference (preferences)	Jordan & Dunham (2021); Moty & Rhodes (2021)
Behavioral *versus* Linguistic (labels, generics, specifics, quantifiers)	Visual (clothes style)	Similarity inference (preferences) Essentialist beliefs	Hoicka et al. (2021)
Linguistic (labels & generics *versus* labels & specifics)	Visual (clothes style)	Essentialist beliefs	Foster-Hanson et al. (2019); Foster-Hanson & Rhodes (2020); Hoicka et al. (2021); Leshin et al., (2021); Noyes & Keil (2020); Rhodes, Leslie, Saunders, et al. (2018); Rhodes et al. (2012)
Linguistic (labels, generics)	Visual (clothes style)	Deontic inference (conforming)	Foster-Hanson et al. (2021) Roberts, Gelman, et al. (2017)
Visual (clothes style)		Deontic inference (conforming)	Foster-Hanson et al. (2021); Roberts, Gelman, et al. (2017); Roberts, Ho, et al. (2017)
Behavioral	Linguistic (labels, generics, specifics) Visual (clothes & hair colors)	Deontic inference (conforming)	Riggs & Long (2020)
Linguistic (labels)	Visual (clothes colors)	Deontic inference (harmful actions, friendship, saving, helping, moral actions)	Chalik & Rhodes (2018); Chalik & Dunham (2020); Chalik et al. (2014); Jin & Baillargeon (2017); Rhodes (2012, 2014); Rhodes & Chalik (2013)
Visual (clothes colors)		Deontic inference (harm)	Chalik & Rhodes (2014)
Linguistic (labels, generics)		Affective responses	Charlesworth et al. (2020); A. M. Gonzalez et al. (2017); Lane et al. (2020)
Visual (clothes colors & countries)		Affective responses	Skinner et al. (2020)

2.2. Cues to form and learn social categories

This section reviews studies in developmental psychology that tested how children identify and learn about social categories, i.e., on what kind of cues children rely to represent novel social categories. Linguistic cues, such as labels and generic statements, have been considered as powerful markers of social category (Rhodes & Baron, 2019) as they convey information about a social category as a whole (Riggs et al., 2014; Shutts & Kalish, 2021). Another perspective considers that perception plays a crucial role in how children form social categories: As infants notice perceptual differences, they also start to develop a tendency to categorize people based on these visual criteria (Bigler & Liben, 2006; Rhodes & Baron, 2019). A related proposition considers that children not only form social categories based on salient visual differences, but also on the behavioral regularities they observe in their environment[3]. This proposition stems from models of learning, proposing that children detect regularities in their environment and draw inferences from those regularities (Bigler & Liben, 2006; Oakes et al., 2009; Park & Hastie, 1987; S. J. Sherman et al., 2013). According to this perspective, children, as "sophisticated pattern-detectors," are able to track common features and behaviors of social agents and, subsequently identify the boundaries of social categories (Riggs, 2019, p. 67; see also Gopnik, 2012). In the following sections, research that employed and tested each kind of cues is reviewed.

2.2.1. Linguistic cues

While a broad area of research, initiated in the late 80s, investigated how labels affected children's inductive inferences about social category members (for e.g., Gelman et al., 1986; Gelman & Heyman, 1999; Gelman & Markman, 1986; Heyman & Gelman, 1999, 2000a, 2000b), these earlier studies did not assess whether labels would lead children to form and identify social categories. One study in particular addressed this question by testing whether labels could help young infants (19-26 months old) to identify members of social categories (Diesendruck & Deblinger-Tangi, 2014). Specifically, in this study, infants were introduced to pictures of real-world category exemplars (e.g., pictures of various women) either accompanied with an abstract label (e.g., "Look a Tiroli") or not (e.g., "Look at this"). The presence of a label helped 19-months toddlers to identify two people as belonging to the same category.

[3] In fact, very early on, infants are sensitive to the regularities they observe in their environment. For instance, Powell and Spelke (2013) showed that from 7 months, infants expected that group members would behave similarly after observing their common behaviors.

This finding was extended to novel social categories (Rhodes, Leslie, Bianchi, et al., 2018). In their study, Rhodes, Leslie, Bianchi, et al. introduced toddlers (28- to 36-months old) to novel social categories either with a label and specific language ("This Zarpie eats flowers"), with a label and generic language ("Zarpies eat flowers") or without label ("This one eats flowers"). The results showed that infants of 32-months (and older) accurately ascribed characters to the same category when a label was used (either with specific or generic language), but not when only visual cues, without label, were provided (e.g., introducing a category member visually marked by clothes color while stating "This one eats flowers"). In addition, Rhodes, Leslie, Bianchi, et al. (2018) found that younger toddlers, between 28 and 32 months old, were able to select accurately individuals as belonging to the same social category only when the social category was introduced with labels in combination with generic language (e.g., "Zarpies eat flowers"). This study revealed an important developmental shift during the second year of life: From 28 months, children can form novel social categories when labels in combination with generic language are employed, but it is only later, around 34 months, that children can form novel social categories based on labels only. Importantly, while labels mark people as belonging to the same group, generic language convey information about the category as a whole: When provided with generic descriptions of a labeled social category, children assume that the communicated information applies to each and every member of the social category (Foster-Hanson et al., 2019; Riggs et al., 2014). Together, these findings show that labels facilitate social categorization: From 19 months, infants can identify similarly labeled people as belonging to the same existing social category, from 28 months infants can form novel social categories based on labels in combination with generic language, and from around 34 months, labels alone are sufficient for infants to form novel social categories.

Please note that this section focused only on studies that investigated how linguistic cues serve the *formation* of novel social categories. On the other hand, an important number of studies investigated the effect of linguistic cues on children's *inferences* about novel social categories. Works inscribed within this latter line of research are presented in Section 2.3, which reviews studies that assessed children's inferences about social categories, and in Section 2.4, which presents in detail how linguistic cues modulate children's inferences about social categories.

2.2.2. Visual cues

Perceptual features are considered as another important cue to form social categories, as they are noticeable with a minimum of effort (Kinzler et al., 2010). According to Bigler and Liben (2006), a marked perceptual distinction between groups of people can be the source of social categorization. This view supports the idea that as soon as children notice a perceptual difference between groups of people, they start to categorize people based on these visual criteria (Aboud, 1988; Baron, 2009; Rhodes & Baron, 2019). However, a contrasting view proposes that, although children notice perceptual similarities among people, these similarities are not sufficient for children to form a rich representation of social categories (Baron et al., 2014; Shutts & Kalish, 2021). Unfortunately, very few studies assessed the formation of social categories based on visual cues only[4]. Most of the time, visual cues are employed as a mean to ensure that children keep track of the social categories during the experimental task. For instance, social categories are visually marked (e.g., by clothing style, hair or clothes colors), while other cues to social categories (e.g., labels, generic statements, behavioral regularity) are the focus of investigation (see Table 1).

Only a couple of studies directly assessed whether visual cues alone would be sufficient for children to form novel social categories (as opposed to studies using visual cues in combination with labels and generic claims). Diesendruck and Weiss (2015) investigated whether children separated people into social categories based on labels and/or physical features. To test this, they used two real-world social categories (race and gender), as well as an artificial social category (shirt colors). After being introduced to the social categories, 5-years old children had to assign different individuals to the given social categories. Importantly, the individuals varied in their degree of similarity with the categories' exemplars (e.g., people wearing shirt colors ranging from red to yellow) and were either referred to by a label or not. In the no-label condition, children categorized individuals based on perceptual similarity for all three social categories (race, gender, and shirt colors). In the label condition, however, children's responses were modulated by the type of social category, i.e., the degree to which the social category is essentialized. Specifically, children disregarded perceptual similarity in favor of labels to categorize individuals differing in shirt colors (the least essentialized category). In contrast, for the gender category (the most essentialized category), children

[4] Please note that, while the present review focuses on novel social categories, an important line of research showed that infants form social categories according to race based on visual cues of faces (e.g., Anzures et al., 2009; Ferera et al., 2021; Lee et al., 2017; Quinn et al., 2016).

ignored the labels and based their categorization on perceptual similarity. For the race category, children's categorization was at chance level. Diesendruck and Weiss' (2015) study showed that from 5 years of age, children were able to form categories based on perceptual similarity. However, when pitted against each other, labels prevailed over visual cues for novel social categories.

Similarly, Baron et al. (2014, Experiment 2) tested whether children would form novel social categories based on shared intrinsic (i.e., skin color) or nonintrinsic (i.e., hat colors) perceptual features. Children between 4 and 7 years old learned about novel social category members (introduced with individuating names) that shared a visual property. This experiment revealed that only 7-year-olds, but not 4-year-olds, were able to predict that characters sharing a visual feature, be it skin or hat color, would behave similarly. Interestingly, in an additional experiment, Baron et al. (2014, Experiment 3) found that children of both age groups were able to predict that social category members would behave similarly when provided with linguistic cues in addition to visual cues. These findings suggest that only 7-year-olds were able to identify novel social categories based on visual cues; younger children needed an additional linguistic cue to form novel social categories. Interestingly, Baron et al. (2014, Experiment 1) also showed that 4-year-olds were able to predict, in the absence of visual cue, that characters sharing the same label would behave similarly. This finding goes in the same direction as the studies of Diesendruck & Deblinger-Tangi (2014) and Rhodes, Leslie, Bianchi, et al. (2018). Altogether, these studies show that children of 5 years and more can form novel social categories based exclusively on visual cues; younger children rely on linguistic cues.

While research investigating children's social categorization based on perceptual similarity is scarce, many studies were conducted with animal categories. An important line of research showed that preschoolers can separate the animal world into categories based on visual similarity or shared internal properties (see for e.g., Diesendruck et al., 1998; Diesendruck & Peretz, 2013). Previous studies also showed that preschoolers expect animals with similar appearances only, i.e., not referred to by labels, to share the same morphological (e.g., "have four stomachs"; Davidson & Gelman, 1990; Sloutsky et al., 2007) and behavioral properties (e.g., "living in nests"; Gelman & Coley, 1990; Jaswal & Markman, 2007). Interestingly, when labels conflicted with visual similarity, children's animal *categorization* was only slightly affected by labels (Diesendruck & Peretz, 2013; Sloutsky & Fisher, 2004), but when children had to predict animals' behaviors, labels guided children's *inferences* (Jaswal & Markman, 2007). These findings on animal categorization contrast with research on social categorization.

While preschoolers tend to readily form novel animal categories from visual similarity, they are less likely to do so for novel social categories, for which they need additional linguistic cues. According to Diesendruck and Weiss (2015), this difference comes from the fact that animal categories, even if they are novel, are essentialized to a greater extent than novel social categories. As such, the degree to which a category is essentialized seems to mediate whether children rely on visual or linguistic cues to determine category membership.

2.2.3. Behavioral cues

Some models of learning propose that children detect regularities in their environment and draw inferences from them (Bigler & Liben, 2006; Oakes et al., 2009; Park & Hastie, 1987; Rhodes & Baron, 2019; S. J. Sherman et al., 2013). According to this perspective, children would infer that some people belong to the same category based on the behavioral regularities they observe in their environment. Many past studies on children's capacity to detect regularities and infer patterns from regularities focused on language learning (e.g., word segmentation, Saffran et al., 1996; or phonetic learning, Maye et al., 2008) or causal learning (e.g., Buchsbaum et al., 2011; Gopnik et al., 2001). Recently, research started to explore how children's capacity to detect regularities plays a role in how they learn and reason about their social environment, and more specifically about social categories. This section reviews recent studies that assessed how observing regular behaviors influenced children's inferences about novel social categories (see Section 2.3 for a detailed review on social category inferences).

Even though the following studies did not directly assess children's formation of novel social categories, we might argue that, if children are able to draw inferences about a social category, they must have identified and acquired a representation of the social category and its members beforehand. This choice of review is motivated by the fact that studies started to employ behavioral cues only recently and are consequently not numerous. For instance, Riggs (2019) investigated whether 4- to 8-years old children would generalize a behavior to social category members based on statistical evidence, i.e., whether a behavior was demonstrated by a large group of people, a subset of people, or a specific individual. Children observed the behaviors of five individuals, belonging either to the same social category or to different social categories. The number of individuals displaying the same behavior (e.g., kneeling before leaving a room) varied across conditions (for details about the conditions, please consult Riggs, 2019, p. 69). The results showed that children consistently inferred the behavior of a novel individual when they observed four category members displaying the same behavior. Children

between 4 and 8 years old can thus attend to behavioral regularities and use them to identify and reason about social categories.

A related study investigated whether representing a high proportion of category members behaving similarly would lead children to generalize the behavior to other social category members (Hoicka et al., 2021, Experiment 2). In this study, 5- and 6-year-olds learned about three members of a novel labeled social category. All three members demonstrated the same behavior (e.g., bouncing a ball on their head). The authors found that children did not generalize a behavior based on their observation of three category members. However, when children were provided with additional linguistic cues to high proportions (i.e., generic claims and proportion quantifiers, such as "most," "many," or "some"), they expected that other social category members would demonstrate the same behavior. It would thus seem that 5-6-years old children have difficulty reasoning about social category members solely based on behavioral cues.

The current state of research using behavioral cues reached contrasting conclusions: While Riggs (2019) found that children could use statistical evidence to reason about social categories, Hoicka et al. (2021) showed that children needed additional linguistic cues, in combination with statistical evidence, to reason about social categories. Three possible explanations might be proposed to account for this discrepancy. First, the number of instances presenting the behavior of category members differed between the two studies. Riggs (2019) presented four category members preforming the same behavior, whereas Hoicka et al. (2021) presented three instances of similarly behaving category members. Second, whereas Riggs' (2019) stimuli consisted of behaviors only (e.g., kneeling, clapping, raising arms, doing push-ups, etc.), the stimuli used in Hoicka et al.'s (2021) study comprised preferences (e.g., love to eat flowers, like to sing, etc.), morphological properties (e.g., have freckles on their feet, have stripes in their hair), and behaviors (e.g., bounce a ball on their head, climb tall fences, etc.). These discrepancies in the type of properties used might have affected whether children drew inferences about category members (see Section 2.3.1 for details on how the type of properties affects children's category inferences). It is consequently possible that the diversity of properties used in Hoicka et al.'s (2021) study in comparison to Riggs' (2019) study impacted on the overall results. Third, the age range of children in Riggs' study was broader than in Hoicka et al.'s study. There might thus be a developmental shift in children's capacity to learn and reason about social categories from observed behavioral regularities that occurs after 5-6 years old. Future studies could assess this possibility (see Section 2.4 for further discussion).

2.3. Inferences about social categories

What do children assume about a person once they know to which social category she belongs? One perspective proposes that predictions about similarity is the main inference that one can draw from social categories (Rhodes & Baron, 2019; Shutts & Kalish, 2021). From this point of view, children assume that social category members share specific properties, such as traits, activities, behaviors, physical attributes, or preferences (Liberman et al., 2017; Over & McCall, 2018). However, simple inferences about shared similarity might be too broad. Research on essentialism extends this perspective: Children might infer that there are deep and stable causal properties determining social category membership (Bigler & Liben, 2006; Foster-Hanson & Rhodes, 2020; Gelman, 2003). On this account, social category members are similar because of an underlying essence, i.e., because they share deep, stable, and intrinsic properties (for e.g., blood composition or internal psychological attributes; Diesendruck & Eldror, 2011; Diesendruck & Weiss, 2015). Importantly, although children can show essentialist thinking from 4 years old, they do not automatically apply essentialist beliefs to novel social categories (Leshin et al., 2021). Research using novel social categories showed that children assume quite easily that category members share observable properties (e.g., behaviors, activities, external physical attributes; see Section 2.3.1), but they do not necessarily hold essentialist beliefs about novel social categories (Hoicka et al., 2021; Vasilyeva et al., 2018). As a result of inductive inferences, children are prone to expect that social category members share observable similarities. With additional inputs (reviewed in Section 2.3.2), children might also come to assume that these observable similarities are caused by an underlying essence.

Moreover, social categories also entail expectations about how social category members ought to behave (Rhodes & Baron, 2019; Shutts & Kalish, 2021). That is, social categories provide "a shared standard against which one's own and another's behavior can be assessed." (Shutts & Kalish, 2021, p. 351) This "shared standard" can concern social norms (e.g., conforming to the regular behaviors of the social category) or social obligations (e.g., engaging in prosocial actions with members of one's own social category) (Bigler & Liben, 2006; Liberman et al., 2017; Over & McCall, 2018; Rhodes, 2013). As a consequence, children tend to ascribe a normative stance to social categories.

Finally, children might also show affective responses toward social category members, and thus develop positive or negative attitudes toward a social category (Charlesworth et al., 2020; Over & McCall, 2018; S. J. Sherman et al., 2013). Past research mainly focused on the

effect of minimal grouping, by showing that simply assigning children to an arbitrary group (e.g., "the Greens") led them to prefer and favor their own group over another group (Kinzler et al., 2010; Park & Judd, 2005; Skinner et al., 2020). However, recently studies started to explore the possibility that children develop affective responses toward a social category without identifying themselves to the social group. This trend of research revealed that children develop affective responses toward a social category by simply observing adults' nonverbal behaviors or by (*over*-)hearing adults' testimonies (Brey & Pauker, 2019; Eggleston et al., 2021; Over & McCall, 2018; S. J. Sherman et al., 2013; Shutts & Kalish, 2021; Skinner et al., 2020).

The following sections present the different types of expectations that children draw about people based on their membership to a novel social category, namely expectations about similarity (Section 2.3.1), essentialist beliefs (Section 2.3.2), deontic inferences (Section 2.3.3), as well as children's affective responses to social category members (Section 2.3.4).

2.3.1. Similarity inferences

Expectation about similarity is one of the main inferences one can draw from social categorization (Rhodes & Baron, 2019; Shutts & Kalish, 2021). For instance, one might expect that a person will have a particular property because this property is characteristic of her social category. The kinds of properties that social categories are expected to share can be diverse and include traits, activities, behaviors, physical attributes, and preferences (Liberman et al., 2017; Over & McCall, 2018; S. J. Sherman et al., 2013). Studies typically assessed whether children would generalize a property they learned about a social category to other individuals belonging to this social category. The present section focuses on similarity inferences that children make about novel social categories (for a review on children's inferences about real-world categories, see Charlesworth & Banaji, in press).

Many studies using novel social categories assessed whether children would expect that social category members share preferences. For example, Jordan and Dunham (2021) found that children between 3 and 9 years old predicted that two individuals, from the same social category, would share activity preferences (e.g., "liking to fly kites"). In this study, social categories were introduced either by a label (e.g., "These kids are called the Zertles") or by similarity (e.g., "These kids like to eat a food called zertles"). In both conditions, children expected that members of the same social category would share activity preferences. Similarly, Moty & Rhodes (2021) found that children, from 4.5 years old, expected that members of the

same social category would be good at a specific activity (e.g., “baking pizzas”). Interestingly, their results also showed that children, after hearing generic claims, made predictions about unmentioned members of an opposite social category; they expected that members of the unmentioned category would not be good at the specific activity (for similar findings, see Kramer et al., 2021). Those results, however, contrast with the findings of Kalish (2012) who showed that 8-years old children did not generalize preferences to social category members. Soley (2019) reached similar conclusions, showing that children between 5 and 7 years old did not expect shared music preference among members of the same labeled social category. However, there is a major difference between these studies. Kalish (2012) and Soley (2019) introduced novel social categories using a label and specific language (e.g., “This Mitu likes this song”), whereas Jordan and Dunham (2021) and Moty and Rhodes (2021) used generic statements in addition to labels (e.g., “Zarpies are good at baking pizzas”). The type of linguistic cues employed to introduce a novel social category thus seems to mediate whether children expect social category members to share preferences (this question is further discussed in Section 2.4).

Another source of variation can come from the nature of the property itself. Recently, Doan et al. (2021) found that children between 2 and 3 years old expected some preferences to be shared, such as food and activities, but that other preferences, like color preferences, are more subjective. Thus, depending on the nature of a property, children will expect it to be generalizable even without generic statement. Riggs et al., (2014) assessed how the nature of a property would influence children’s expectations. They introduced 3- to 5-year-olds to a member of a novel labeled social category with her associated property. The nature of the property varied across conditions, and was either morphological (e.g., “has plasma in the blood”) or historical (e.g., “got bit by a krem in the meadow”). Children generalized more morphological than historical properties to other category members. Similarly, Diesendruck and Eldror (2011) showed that children generalized biological properties (e.g., “have melanin in their blood”) when they were presented as causing external properties (e.g., “have round bones”). In addition, Soley (2019) found that children predicted that social category members would share knowledge about music but would not necessarily like the same music. Therefore, children might consider activities, food preferences, physical features, and cultural knowledge as more generalizable than other kinds of properties.

What about psychological properties, such as traits? Relatively little research assessed whether children extend traits to novel category members despite theoretical approaches of

stereotyping arguing that people tend to apply traits to social categories (e.g., women are caring) (Bigler & Liben, 2006; S. J. Sherman et al., 2013). Would children predict that a person is generous because she belongs to a social category that is characterized as generous? Although this question has not yet been directly addressed, studies on trait attribution provide possible answers and interesting perspectives. For example, Liu et al. (2007) showed that from 4 years of age, children are able to infer traits (e.g., selfish) from behaviors (e.g., not sharing), and vice versa. In addition, Heyman and Gelman (2000b) found that children expected that individuals described with the same trait (e.g., shy/outgoing) would share preferences. As such, children can attribute traits to people and make inductive inferences from traits: If two people are shy, they will engage in the same kind of activities (Heyman & Gelman, 2000a). In these previous studies, psychological traits were used to introduce a category (e.g., "shy people"). That is, children viewed traits as categories (C. M. Gonzalez et al., 2010). From these findings, it is most likely that children would also generalize psychological traits based on social category membership.

To summarize, the studies reviewed here revealed that children's similarity inferences are mediated by (a) the nature of the property, namely if the property is viewed as generalizable, children will expect it to be shared among category members marked by labels, and (b) the type of linguistic cue employed, specifically using generic claims or high proportion quantifiers will prompt children to assume that any property applies to every member of a social category.

2.3.2. Essentialist beliefs

Children assume quite easily that some properties are generalizable across category members, but they do not necessarily believe that there is an underlying essence that causes category members to be similar. In fact, expecting that a property is shared among social category members does not imply holding essentialist beliefs about the social category (Hoicka et al., 2021; Vasilyeva et al., 2018). It is only with additional inputs that children may develop essentialist beliefs about social categories and view them as natural kinds (Rhodes, 2013). Essentialist beliefs about social categories consist in inferring that there are stable, natural, and intrinsic causes to social category membership (e.g., boys are good at math because of some inherent properties specific to boys) (Bigler & Liben, 2006; Gelman, 2003; Rhodes & Baron, 2019; Shutts & Kalish, 2021). This belief also leads people to view social categories as discrete, such that members of two social categories are fundamentally different and members of the same social category are deeply similar (Foster-Hanson & Rhodes, 2020; Peretz-Lange &

Muentener, 2020). These essentialist assumptions seem to emerge under certain conditions, where emerging cognitive biases and specific cultural inputs play an important role (Leshin et al., 2021; Rhodes, 2013; Rhodes et al., 2012). This section reviews studies that investigated whether children develop essentialist beliefs about novel social categories and if so, under what conditions (for a review of essentialist beliefs about existing social categories, see Peretz-Lange & Muentener, 2020).

To assess children's essentialist beliefs, researchers typically described the property of a novel social category, and then asked children to explain what caused the property, whether the property was inherited, and whether the property is generalizable to other category members. Regarding causality inferences, children's explanations about what caused the category property are assessed with open-ended questions (e.g., "Why does this person do X?"). If children mention social category membership as being the cause of the property (e.g., "Because she is a Zarpie"), their answer is considered as essentialist. Multiple studies showed that simply using generic language when presenting a novel social category and its property (e.g., "Zarpies sleep in tall trees") led 4- to 9-years old children to provide more explanations of the property as caused by social category membership than when they learned about social categories' properties through specific language (Cimpian & Markman, 2011; Leshin et al., 2021; Rhodes et al., 2012; Rhodes, Leslie, Saunders, et al., 2018). This was also true when high proportion linguistic cues were employed, such as "most" and "many" (Hoicka et al., 2021), and when a generic claim was subsequently negated (e.g., "That's not right about Zarpies," Foster-Hanson et al., 2019). Interestingly, Foster-Hanson and Rhodes (2020) reached similar findings when the novel social categories were presented as created by a powerful being. Specifically, when the category was described as created by a powerful being (instead of being discovered by a standard person), 4-5-year-olds explained more often that the reason why a person had a certain property was due to her category membership. These studies showed that generic claims, high proportion linguistic cues, and the type of explanation provided about the category's origin led children to infer category membership as the cause of the property's occurrence.

Essentialist beliefs are also assessed with switched-at-birth paradigms to test children's inferences about the inheritance of a property. The task consists in asking children whether a member of social category A, raised in a family of social category B, will have the properties of the birth parents (from social category A) or of the adoptive parents (from social category B). If children expect that the individual inherited the property of category A, the answer is

considered as essentialist. A first influential study showed that 4-year-olds and adults inferred more often that a property would be inherited when hearing generic compared to specific statements (Rhodes et al., 2012). However, subsequent studies revealed that generic claims and the origin of social category did not lead children (between 4 and 7 years old) to expect that the properties would be inherited (Foster-Hanson et al., 2019; Foster-Hanson & Rhodes, 2020; Hoicka et al., 2021). Interestingly, Noyes and Keil (2020) recently showed that the type of properties influenced children's responses in a switch-at-birth paradigm. When presented with generic statements, children from 6 years old – but not younger children – expected more that biological properties would be inherited (e.g., "feeling sick when drinking milk") than cultural properties (e.g., "believing that fish talk to God") (Noyes & Keil, 2020). These different patterns of results might be explained by a developmental change occurring between 4 and 9 years of age as suggested by Leshin et al.'s (2021) and Noyes and Keil's (2020) studies, which revealed that children make more refined inferences about properties' inheritance as they grow up. Consequently, the conditions under which children draw inheritance inferences are subject to developmental changes and remain to be specified.

Finally, essentialist beliefs have also been measured with induction questions: Would children infer that A shares properties with B because both are members of the same social category? As previously argued, expecting that a property is shared among category members does not necessarily imply that there is an underlying essence causing category members to be similar (Hoicka et al., 2021; Vasilyeva et al., 2018). For this reason, some researchers assessed essentialist beliefs by other means, such as questions pertaining to the inflexible aspect of a property (e.g., "Would this person only do X or would she sometimes do Y?"; Leshin et al., 2021) or to the extent to which the property is generalizable (e.g., "How many other members will do X?"; Foster-Hanson et al., 2019). Results showed that for each of these assessments, children provided answers congruent with essentialist beliefs when they learned about social categories through generic language. Specifically, 4-9-year-olds expected the property to be inflexible and absolute when they learned about the property through generic language (Leshin et al., 2021). And 5-7-year-olds generalized a property more broadly – to more social category members – when generic claims were used compared to specific statements (Foster-Hanson et al., 2019). Together, these findings provide evidence that generic claims prompted children to infer that a property would be inflexible and broadly generalizable.

2.3.3. Deontic inferences

Children do not only expect similarity among category members, they also hold expectations about social obligations. Deontic inferences can relate to how social category members ought to behave (e.g., going to church) and to how they should interact with each other (e.g., helping each other). In fact, children tend to ascribe a normative stance to social category members' behaviors, namely, they consider that not conforming to one's own social category is unacceptable (Foster-Hanson et al., 2021; Roberts, Gelman, et al., 2017). In addition, social categorization supports inferences about how people are obligated to one another (Rhodes, 2013), for instance that members of the same social category are expected to cooperate (Kaufmann & Clément, 2014), restrain from harming each other (Rhodes & Chalik, 2013), and engage in prosocial behaviors (Liberman et al., 2017). Detecting social groupings has an important role in determining patterns of cooperation and competition (Kinzler et al., 2010; Rhodes & Baron, 2019). Similarly, adherence to norms and preference for conformity are proposed to support cultural learning and coordination throughout evolution (Chudek & Henrich, 2011; Roberts, Gelman, et al., 2017). Therefore, based on social category membership, children might rapidly form expectations about social obligations, namely that individuals should conform to their social category and engage in prosocial behaviors with fellow category members.

Children, from a very young age, hold expectations about interpersonal interactions between category members, even when social categories are novel. The first line of studies exploring children's predictions about how category members should interact focused on harmful actions. From 3 years old, children were more likely to expect that harmful actions would be directed toward members of different social categories than toward members of the same social category (Chalik et al., 2014; Chalik & Rhodes, 2014; Rhodes, 2012). Other studies found that by age 4, children explained that harmful actions were perpetrated because of category membership (Chalik & Rhodes, 2015; Rhodes, 2014). In addition, children of 4.5 years old viewed harmful actions as wrong when they involved people of the same social category, but not when the harmful actions were directed toward members of a different social category (Rhodes & Chalik, 2013). Children's evaluations remained the same even when a rule stated that it was permissible to commit harmful actions (Rhodes & Chalik, 2013). These results attest that children expected members of novel social categories, introduced by linguistic (labels) and visual (shirt colors) cues, to be intrinsically obligated not to harm each other.

Do children have expectations regarding only harmful actions, or do they also hold expectations about other types of intragroup relations? Chalik and Rhodes (2018) investigated whether children's expectations about category members extended beyond harmful actions to include other social obligations. In their study, Chalik and Rhodes tested (a) harmful behaviors (e.g., stealing, hitting), (b) prosocial actions (e.g., sharing, hugging, helping), (c) friendship relations, and (d) saving behaviors. Similar to the above-mentioned studies, preschool children predicted above chance that harmful actions would be directed at the members of another social category. Regarding the other predictions, children, around 4 years old, expected members of the same social category to engage in friendship relations and saving behaviors, but not consistently in prosocial actions. This latter result should be clarified in the light of Jin and Baillargeon's (2017) findings, who showed that 17-month-olds expected members of the same social category to help each other. These contradictory results might be explained by the difference in the type of prosocial behaviors tested: Whereas Jin and Baillargeon only included helping behaviors, Chalik and Rhodes (2018) tested multiple prosocial behaviors (such as hugging, sharing, and helping). Finally, a study conducted by Chalik and Dunham (2020) revealed that expectations relative to social obligations were mediated by morality. In fact, Chalik and Dunham found that 4- to 5-year-olds expected members of the same social category to demonstrate positive moral actions toward each other and to engage in negative moral actions toward another category. When actions were not morally laden, children provided no specific prediction. Altogether, this line of research revealed that, already from 17 months old, children predicted that members of the same social category should engage in positive moral behaviors.

A related line of research revealed that children expected individuals to conform to their social category. Roberts et al. (2017) investigated whether children drew prescriptive judgements after being exposed to a social category's regular behavior. More precisely, 4- to 13-years old children learned about the regular properties of two novel social categories. The information was provided with generic statements (e.g., "Hibbles eat these kind of berries"). Children were then invited to evaluate the behavior of a new category member, that either conformed to the social category or not. The study revealed that children between 4 and 13 years disapproved more of non-conforming individuals than conforming individuals (for similar findings, see Foster-Hanson et al., 2021). Riggs and Long (2020) provided further evidence that children (4-8 years old) viewed a behavior as a norm, when they observed a majority of members from the same social category demonstrating the same behavior. Interestingly, children in Riggs and Long's study also predicted that a norm implied that all

members of a social category should demonstrate the same behavior. In addition, several studies demonstrated that children, around 4 years, generalized more norms than preferences to social category members (Kalish, 2012; Kalish & Lawson, 2008; Riggs et al., 2014a). That is, children expected that another "Lissian" would conform to her social category when a property was introduced as a norm (e.g., "has to travel by train") but not as a preference (e.g., "likes to travel by train"). Hence, children, from 4 years of age, do not only view norms as prescriptive, but they also draw prescriptive judgements from observing regular behaviors.

But would children also infer a norm solely based on category membership, rather than on the regularity of the behavior? Roberts, Ho, et al. (2017) investigated the kinds of information, on which children of 4 and 9 years based their prescriptive judgments about category members. The authors tested four kinds of cues with which novel social categories were introduced to children: simple group presence (e.g., three individuals of each group was presented in the background), use of a label only (e.g., "this Hibble eats this kind of berries"), use of a label with a generic claim (e.g., "Hibbles eat this kind of berries"), or without any particular information (e.g., "This one eats this kind of berries"). The experimental design replicated Roberts, Gelman, et al. (2017); children judged if the behavior of an individual was okay or not, depending on whether the individual acted in conformity with her category or not. The results revealed that using simple labels, presenting category members, or uttering generic statements led children to disapprove of the behavior of a non-conforming individual. Roberts, Ho, et al. (2017) nonetheless noted an important difference between younger (4-6-years-olds) and older children (7-9-years-olds): It is only when provided with generic statements that younger children disapproved of non-conformity above chance. From 7 years old, children can rely on non-linguistic cues to make deontic inferences about novel social categories, whereas younger children, in the absence of generic language, have difficulty inferring how social category members ought to behave.

2.3.4. Affective responses

Showing positive or negative affective responses toward a social category corresponds to having positive or negative attitudes, which can subsequently serve as a basis for prejudices or even discrimination (Fiske & Taylor, 2013; Kurdi, Mann, et al., 2019). According to researchers in the field of social identity, the mere act of categorizing people suffices to produce a preference for one's own group over any outgroup (Bigler & Liben, 2006; Over & McCall, 2018; Spelke & Kinzler, 2007; Tajfel, 1970). That is, simply assigning children to an arbitrary

group (e.g., “the Greens”) leads them to prefer and favor their own group over another group (Kinzler et al., 2010; Park & Judd, 2005; Skinner et al., 2020). Studies that induced ingroup and outgroup distinctions with minimal group paradigms are not presented here, as they constitute a broad area of research and would be beyond the scope of the present article (for reviews, see Dunham, 2018; Over & McCall, 2018). However, recent studies revealed that, without identifying themselves with a social group, children developed affective responses toward a social category by simply observing adults’ nonverbal behaviors (Brey & Pauker, 2019; Eggleston et al., 2021; Over & McCall, 2018; S. J. Sherman et al., 2013). In fact, former research with real-world social categories found that children’s affective attitudes were correlated with those of their parents (for a meta-analysis, see Degner & Dalege, 2013). This important conclusion inspired a new area of research interested in determining how one communicates about, or interacts with, a novel social category affects children’s evaluations of this social category. Specifically, children might develop affective responses toward a social category by directly observing adults’ behaviors or by (*over*-)hearing adults’ testimonies (Shutts & Kalish, 2021; Skinner et al., 2020). The present section thus reviews studies that explored how explicit messages and nonverbal behaviors shape children’s affective responses toward novel social categories.

Testimony allows one to acquire information without direct observation (Harris, 2012), and can also serves as a mean to learn about social agents (Dunbar, 2004). Some have thus argued that children can learn how to evaluate social agents based on verbal statements (Charlesworth et al., 2020; Lane et al., 2020). In fact, Skinner and Meltzoff (2019) conducted a systematic review on the factors influencing children’s intergroup biases about real-world categories (such as age, race, weight, gender, disability, or ethnicity). Their review revealed that one factor in particular was associated with increased intergroup bias: negative messages about social categories. Similarly, studies using novel social categories showed that children acquired novel implicit attitudes after hearing detailed stories describing the valenced behaviors of social category members (e.g., “Lups are very nice. They listen to what their parents tell them to do, they clean their room, …”; A. M. Gonzalez et al., 2017), but also after hearing a single evaluative statement about a novel social category (e.g., “Longfaces are good”; Charlesworth et al., 2020). Therefore, children’s implicit attitudes can emerge from overt messages about social categories.

But one might argue that it is unlikely that adults would directly express evaluative statements about social categories. A new area of research started to determine whether children

would also form positive and negative attitudes from overheard messages. Lane et al. (2020) tested this hypothesis by providing 4- to 9-years old children with negative information about a novel social category (e.g., “Those Gearoos are really bad people”). The information was either directly addressed to children or overheard as part of a phone conversation (between the experimenter and a fake interlocutor). Children, who heard negative information, both directly and by eavesdropping, were less willing to be friends with the negatively portrayed social category and evaluated this social category more negatively. These findings represent the first evidence that children are sensitive to overheard negative messages about social categories.

In addition, adults are likely to communicate their own positive and negative attitudes through nonverbal behaviors. Several studies first explored the extent to which positive and negative nonverbal behaviors, demonstrated by an adult toward a specific individual – not associated to a social category – would influence children’s evaluation of the individual (e.g., Brey & Pauker, 2019; Brey & Shutts, 2018; de Rosnay et al., 2006; Skinner et al., 2017). Those former studies revealed that children were sensitive to adults’ nonverbal behaviors when evaluating other people. Furthermore, Skinner et al. (2017) found that 4-5-years old children, after observing an adult’s positive and negative nonverbal behaviors toward two individuals, showed a preference for the target of positive nonverbal behaviors, as well as the target’s closest friend. This finding suggests that children acquired attitudes via nonverbal behaviors and that they generalized those attitudes to other individuals based on friendship relations (for similar findings, see Eggleston et al., 2021). In a subsequent study, Skinner et al. (2020) assessed whether children would generalize their attitudes to other individuals based on social category membership. In this study, preschool-aged children were introduced to two novel social categories and observed a neutral individual acting positively toward one social category member and negatively toward one member of another social category. Children again preferred the individual for which positive nonverbal behaviors were demonstrated, and this preference extended to the whole social category. Skinner et al.'s (2020) study provides evidence that exposure to valenced nonverbal behaviors toward a single category member can lead children to prefer a social category over another. Thus, nonverbal behaviors might play a crucial role in how children develop positive and negative attitudes toward social categories, especially because emotional reactions are pervasive and difficult to control (Clément & Dukes, 2017; Weisbuch & Ambady, 2008). Children might inevitably be exposed to and spot these nonverbal signals in their everyday life, and in turn use them to evaluate social agents.

2.4. Summary and propositions for future studies

Social categorization involves two crucial processes: Children first learn to form social categories and then, based on social category membership, children make inferences about, and develop affective responses toward category members (Baron, 2009; Baron et al., 2014; Over & McCall, 2018; Shutts & Kalish, 2021). This final section articulates how different types of cues to introduce novel social categories influence children's categorization of, and inferences about social categories. Opportunities for future research are also proposed.

2.4.1. Going beyond linguistic cues

From the research reviewed in this article, it seems that linguistic cues prevail over visual cues both on children's learning of, and inferences about novel social categories. This is somewhat contradictory with the proposition that visual cues should very early serve as a basis for the detection of social categories because they require a minimum of effort (Kinzler et al., 2010). These findings also contradict research in the animal domain showing that, without label, children tend to categorize animals and make inferences about animals based on their similar appearances (see for e.g., Davidson & Gelman, 1990; Diesendruck et al., 1998; Diesendruck & Peretz, 2013; Gelman & Coley, 1990; Jaswal & Markman, 2007; Sloutsky et al., 2007). A possible explanation might be that most of the visual cues employed in the studies reviewed here were too artificial to signal social category membership, as they consisted mostly of clothing style or clothing colors (except for Baron et al.'s (2014) study that used intrinsic properties, namely skin color, see Table 1). In fact, from an evolutionary perspective, other cues would be more informative than clothing, such as language, ritualistic behaviors, psychological traits, and group entitativity (Liberman et al., 2017; Liu et al., 2007; S. J. Sherman et al., 2013). Moreover, past research suggests that the degree to which a category is essentialized mediates whether children rely on visual cues to determine category membership (Diesendruck & Weiss, 2015). To date, visual cues were mainly employed to help children keep track of social categories during the experimental manipulation and were not necessarily considered *per se* as cues on which children would rely to identify and make inferences about novel social categories. Future studies could thus vary the extent to which novel social categories are essentialized and employ more informative perceptual cues alone and in combination with linguistic cues. This would help to better understand whether linguistic cues constitute indeed a privileged means of learning and making inferences about novel social categories.

Surprisingly, although theoretical accounts propose that children attend to behavioral regularities in their environment to identify social categories (Bigler & Liben, 2006; Oakes et al., 2009; Park & Hastie, 1987; S. J. Sherman et al., 2013), this possibility has rarely been addressed in experimental research. Besides, the current field of research reached divergent conclusions: While Riggs (2019) found that children relied on perceived behavioral regularities to make inferences about novel social categories, Hoicka et al. (2021) showed that children needed additional linguistic cues. There is thus a clear need for research on this topic, and future studies should investigate how observed regularities shape children's representations of social categories. For example, would children assume that, if a majority of individuals demonstrate the same property, these individuals belong to the same social category? Would children need to observe the property in 5, 10, 20 or more social category members? (The actual state of research only tested behaviors that occurred 100% of the time among 3 to 4 category members). What other inferences would children draw from observing regularities? Determining whether children learn about social categories based on the regularities they observe in their environment represents a great opportunity to tackle the possible role of contextual factors on children's acquisition of stereotypes.

2.4.2. Opening ways to study the acquisition of essentialist beliefs

Generic claims and noun labels (e.g., "She is a carrot-eater") have also been documented to play an important role in the development of essentialist beliefs (Gelman & Heyman, 1999; Leshin et al., 2021; Markman, 1989; Rhodes & Baron, 2019; Walton & Banaji, 2004). The studies reviewed in this article revealed that children, from 4 years of age, hold essentialist beliefs about novel social categories when the latter were introduced with generic claims (but not with high proportion quantifiers). Moreover, recent studies showed that children allocated fewer resources to essentialized social categories (i.e., social categories introduced with generic claims) compared to categories that were not essentialized (i.e., introduced with specific language) (Leshin et al., 2021; Rhodes, Leslie, Saunders, et al., 2018). These studies extend previous findings: While previous research using minimal group paradigms showed that children allocate more resources to their in-group in comparison to an out-group (see for e.g., Benozio & Diesendruck, 2015; Dunham et al., 2011; Plötner et al., 2015), Leshin et al.'s (2021) and Rhodes, Leslie, Saunders, et al.'s (2018) studies revealed that children tend to refrain from allocating resources to essentialized categories. Altogether, there is considerable evidence that generic language fosters essentialist beliefs.

Yet, this general conclusion should be nuanced. Essentialist beliefs have been traditionally assessed as part of a composite score that collapses various measures of essentialism, including inheritance, causality, and induction inferences (Leshin et al., 2021). Taken separately, however, essentialist belief measures yielded contrasting results. Whereas generic claims led children to infer a property as inflexible (Leshin et al., 2021), broadly generalizable (Foster-Hanson et al., 2019), and caused by category membership (Cimpian & Markman, 2011; Leshin et al., 2021; Rhodes et al., 2012; Rhodes, Leslie, Saunders, et al., 2018), generics did not necessarily lead children to assume that a property was inherited (Foster-Hanson et al., 2019; Hoicka et al., 2021). Please note however, that school-aged children assumed that *biological* properties were inherited when presented with generic descriptions (Noyes & Keil, 2020). The current state of research thus seems to suggest that generics foster some types of essentialist beliefs rather than essentialist thinking in general. Future research on essentialist beliefs could clarify this possibility.

Multiple studies also showed that specific cultural factors can be responsible for the development of social essentialism. Until recently, generic language was predominantly studied as a way by which children develop essentialist beliefs. In fact, generic language constitutes a privileged cultural input that guides children's acquisition of essentialist beliefs, not only because they are understood as "communicating nonaccidental generalizations," (Rhodes et al., 2012, p. 13527) but also because they appear in 3-4% of parent-child exchanges (Gelman et al., 2008). However, a new area of research started to propose that other factors can foster social essentialism, such as explanations pertaining to the origin of social categories (Foster-Hanson & Rhodes, 2020), information regarding the social category's deep and intrinsic properties (Diesendruck & Weiss, 2015), personal motivational factors, such as the need to belong (Diesendruck, 2021), or the existence of segregation and social hierarchies (Rhodes et al., 2012; Rhodes, Leslie, Saunders, et al., 2018). Crucially, social essentialism can have serious implications: People holding essentialist beliefs assume that differences between social categories are due to some intrinsic and inflexible causes rather than to external and contextual factors (Cimpian & Salomon, 2014; Rhodes, Leslie, Saunders, et al., 2018; Salomon & Cimpian, 2014). As such, people who apply essentialist beliefs to social categories are more likely to accept social differences and inequalities and tend to disregard information about an individual (Cimpian & Salomon, 2014; Foster-Hanson & Rhodes, 2020; Rhodes, Leslie, Saunders, et al., 2018; Shutts & Kalish, 2021). In an endeavor to better understand the origin

of negative intergroup biases, future studies should thus investigate under which circumstances children develop essentialist beliefs.

2.4.3. Exploring the development of affective responses toward social categories

Recent studies explored how children ascribe affective responses to social categories. Studies assessing the impact of direct or overheard messages introduced novel social categories with linguistic cues, namely labels and generic language (for e.g., "Longfaces are bad"; Charlesworth et al., 2020; A. M. Gonzalez et al., 2017; Lane et al., 2020). To date, only one study assessed the effect of positive and negative nonverbal behaviors on children's evaluations of novel social categories, marked with minimal visual cues (Skinner et al., 2020). This latter study suggests that children might generalize affective responses to social category members even when their category membership is signaled with minimal cues.

The studies reviewed here focused mainly on the effects of others' emotions and testimony on children's development of affective responses toward novel social categories (Charlesworth et al., 2020; A. M. Gonzalez et al., 2017; Lane et al., 2020; Skinner et al., 2020). Yet, affective responses toward a social category can also be related to the kinds of properties that are associated with a social category (e.g., disliking a social category because it is thought to be lazy) (Shutts & Kalish, 2021). As argued in the Introduction, associating properties to social category members corresponds to forming stereotypes. If those properties are viewed as positive or negative, then stereotypes take on an affective aspect and become a source of positive or negative attitudes toward a social category (Allport, 1954). For instance, it is possible that a child who learns that a social category is generous, would more likely prefer and engage in friendship relations with this social category than with a social category depicted as stinky. Research on intergroup bias could thus benefit from future research on how valenced properties, and especially valenced psychological traits, might impact children's affective responses toward novel social categories.

2.5. Conclusion

The present article aimed to present the actual state of research in developmental psychology that investigates how children learn and make inferences about novel social categories. By delineating how specific cues to introduce social categories lead to different

types of inferences, this article provides some key insights on the way children might develop stereotypes, positive and negative attitudes, as well as intergroup biases. In the future, researchers should keep on using novel social categories in their experimental designs, because using novel social categories has the advantage of shedding light on general principles of social categorization, that are not tied to a specific cultural background. Studies with novel social categories offer a unique opportunity to compare and replicate findings across cultural contexts and to possibly identify universal conditions under which children learn and make inferences about social categories.

Summary and comments

Chapter 1 provided an overview of recent research in developmental psychology that assessed how children learn to make inferences about novel social categories. Three types of cues have been put forward as means to identify social categories, namely linguistic, visual, and behavioral cues. Then, as a result of inductive inferences, children expect that social category members will share specific properties. While this review shed light on the kinds of inferences children make about novel social categories, a few important questions remain open.

First, children seem to conceive some properties as more generalizable than others, but why is it so? One perspective holds that children need sufficient world knowledge to consider whether a property is generalizable or not (Riggs et al., 2014a). For instance, empirical works showed that children from around 5 years of age have an understanding of the shared nature of internal biological features (e.g., Gelman, 2003) and deontic properties (e.g., Kalish & Lawson, 2008). However, it could also be argued that the kind of properties children will consider as generalizable might be directly linked to the nature of the social category. For instance, a category that is biologically defined might lead to the generalization of biological properties, a category based on religion might lead to generalize habits and customs, or a category defined by occupational roles might lead to generalize skills and obligations (e.g., Noyes et al., 2021). An avenue for future research would therefore be to systematically assess what properties are generalized depending on the defining features of a given social category, i.e., how the nature of a social category licenses different domains of inferences.

Second, while investigating categorization processes with novel social categories might reveal a valuable tool to assess general aspects of social categorization, it also suffers from important drawbacks. Indeed, using novel social categories may limit the extent to which

children emotionally engage with the categories and they might not show the same motivational stance toward novel categories as they would toward real ones. This can limit the generalizability of findings and the eco-logical validity of studies employing novel social categories. The main challenge when resorting to fictional worlds in research is that we do not currently know for sure whether children's reasoning in the pretend-mode accurately corresponds to real-word situations (cf. Rakoczy, 2022)

3. Chapter 2

Developmental changes in category-based inductions: The effects of labels and statistical evidence on children's inferences about novel social categories[5]

Abstract

The psychological mechanisms that subserve inductions about novel social categories in childhood are hotly debated. While research demonstrated that language, and in particular generic statements, plays a major role in how children learn to attribute properties to social categories, developmental theories propose other mechanisms. One theoretical account holds that the mere act of labeling social categories is sufficient for children to generalize properties to category members, because labels are considered as referring to significant, homogeneous kinds of people. A second theoretical account proposes that children generalize properties to category members from statistical evidence, i.e., by directly observing regularities and correlations between category members and certain properties. The present study assessed those two hypotheses, by testing (via an online experiment) the effects of simple category labels and observation of statistical evidence on European 4- to 6-year-olds ($N = 88$) and 7- to 9-year-olds' ($N = 92$) predictions about novel social categories. From around 7-9 years, children generalized properties to category members based on simple category labels or on their observation of a majority of unlabeled category members having the same property. Four- to 6-year-old children, however, made similarity inferences only when both labels and statistical evidence were combined. Overall, the present study highlights a developmental shift from an early limited tendency to make similarity inferences to a later propensity to infer similarity from small evidence. These findings deepen our understanding of the conditions under which children start to make similarity inferences. Implications for the acquisition of stereotypes are also discussed.

[5] Reprint of: Mari, M. A., Clément, F., & Paulus, M. (under review). Developmental changes in category-based inductions: The effects of labels and statistical evidence on children's inferences about novel social categories. *Developmental Psychology*.

This research was funded by the Swiss National Science Foundation, under the grant number P1NEP1_191722.

3.1. Introduction

Categories are extremely useful cognitive tools for making inferences about the world, notably the social environment. Based on social category membership, children, like adults, can predict and explain new individuals' actions by drawing on past experiences (Baron, 2009; Rhodes, 2013; Rhodes & Baron, 2019). For example, if playing the piano is a known feature of social category A, then *x*, as a member of category A, is likely to also play the piano. Traditionally, two processes are involved in social categorization: First, children seek culturally relevant properties on which they can categorize individuals, i.e., they learn to form social categories (Baron et al., 2014; Mari, 2022; Over & McCall, 2018; Shutts & Kalish, 2021). For example, in some cultural contexts, religion may be more salient to categorize people, whereas in other contexts, categorization may be based on race or social class (Segall et al., 2015). Second, children may learn to associate certain properties (e.g., traits, roles, activities, behaviors, physical attributes, or preferences) with specific social categories (Devine, 1989; Liberman et al., 2017; Over & McCall, 2018; Shutts & Kalish, 2021). Children then make similarity inferences, expecting that members of the same social category share similar properties (Heyman & Gelman, 2000a, 2000b; Rhodes & Baron, 2019; Shutts & Kalish, 2021). Referring to the above example, children may expect that a person is good at playing the piano because they have learned that this property is characteristic of the person's social category. In the current paper, we focused on how children learn to *associate* properties with social categories, rather than on how they learn to *form* social categories.

Importantly, attributing properties to social category members is a key process in the formation of stereotypes. Stereotypes correspond to sets of expectancies about the properties of social categories, and are shared to a greater or lesser extent among people from the same community[6] (O. Klein & Bernard, 2015). On this account, social categorization and stereotyping are closely related: As one learns to associate some properties with a social category (e.g., women are caring, men are good at math), one constructs a set of stereotypical expectancies about the members of this social category (Dovidio, 1999). Therefore, a better understanding of how children generalize properties to social categories can provide key insights on the origin of stereotypes. In the following, we present two theoretical perspectives

[6] Please note that stereotypes, unlike attitudes and prejudices, do not necessarily involve a positive or negative evaluation of a social category (e.g., one can hold a stereotype that homosexual couples have aesthetic flair vs. believing that homosexual couples are deviant) (Fiske & Taylor, 2013; Kurdi, Mann, et al., 2019).

of how children learn to attribute properties to novel social categories, namely from minimal linguistic markers and from observation of statistical evidence.

A highly efficient way by which children can acquire knowledge about the social environment is through language (Gelman & Roberts, 2017; Harris, 2012). Theories of stereotyping argue that stereotype acquisition might rely on how cultural environments structure the social world by means of language, especially by labeling certain social groups (Bigler & Liben, 2006, 2007; Hirschfeld, 1996; Shutts & Kalish, 2021). Noun labels referring to categories (e.g., "she is *a carrot-eater*", "John is *an intellectual*") have been shown to be particularly important in shaping children's inferences about the stability of a person's characteristics in comparison to verbal predicates (e.g., "she eats carrots whenever she can") or adjectives (e.g., "John is intellectual") (Gelman & Heyman, 1999; Markman & Smith, described in Markman, 1989; Walton & Banaji, 2004). Furthermore, category labels are interpreted by children as marking social divisions (Over & McCall, 2018) and as referring to significant kinds, about which there is likely more to learn (Shutts & Kalish, 2021; Waxman & Markow, 1995). For instance, upon hearing sentences such as "Girls go to the right, boys to the left" or "The Zarpies go first, then the Gorps," children might understand that there is something generally important about this social division and use it to make category-based inferences (Bigler & Liben, 2007; S. J. Sherman et al., 2013; Shutts & Kalish, 2021). Category labels are thus powerful markers of social categories (Rhodes & Baron, 2019), and might have a privileged effect on how children learn about the properties of social categories and develop stereotypical expectancies.

A broad area of research has shown that category labels may license strong inferences and essentialist beliefs about social categories when combined with generic statements (e.g., "Zarpies eat flowers", see for e.g., Baron et al., 2014; Diesendruck & Deblinger-Tangi, 2014; Foster-Hanson et al., 2019; Gelman & Markman, 1986; Jordan & Dunham, 2021; Leshin et al., 2021; Rhodes et al., 2012, 2018; Waxman & Markow, 1995). Generic language constitutes a particular form of language, conveying information about the category *as a whole* (Foster-Hanson et al., 2019; Riggs et al., 2014a) and implying that the information is essential to the identity of its referents (Bloom, 2004; Cimpian & Markman, 2011; Gelman, 2003). As such, when children are provided with information about a labeled category in the form of generic statements (e.g., "Boys are good at math," "Zarpies play the piano"), they are prone to infer that the speaker intended to provide meaningful and relevant information that applies to *all members* of a specific *kind* of people (Foster-Hanson et al., 2019; Mari, 2022).

But is it necessary for children to hear generic statements about social categories to make similarity inferences? A few previous studies showed that children generalize properties to people that are referred to by the same label only (without generic statement), e.g., "This boy is *Jewish*," "This one is a *Wayshan*" (Diesendruck & haLevi, 2006; Diesendruck & Weiss, 2015; Kalish, 2012; Waxman, 2010). For example, Diesendruck and haLevi (2006) provided evidence that category labels are a powerful source of information for similarity inferences. Across a series of experiments, 4- to 6-years old children were asked to predict the properties of a target character, after being introduced to two other characters that shared personality traits (e.g., "being shy") or category membership (e.g., "being Jewish") with the target. When pitted against personality traits, children were more likely to use category labels to predict the behavior of the target character (e.g., they were more likely to predict that Jewish characters, but not shy characters, will play "zaber"). In other words, children of 4-6 years old predicted that characters from the same social category (e.g., Jewish), but not with the same personality trait (e.g., shy), would share the same properties (see also Diesendruck & Weiss, 2015). Waxman (2010) further assessed the role of labels on category-based inferences. Across two experiments, children of 4 years old were introduced to individuals (e.g., a Black woman, a White woman) and their associated properties (e.g., "is good at a new game called zaggit"). In one condition, the individual was introduced with a label (e.g., "This one is a *Wayshan*"); whereas in the other condition, no label was used (e.g., "This one *eats big lunches*"). In the no-label condition, children projected the property across people in general, without taking into account race or gender. However, when individuals received arbitrary category labels, children predicted that only members of the same race (Experiment 2) or same gender (Experiment 3) would share the property. These findings provide evidence that children are sensitive to category labels and can use them to generalize properties to social category members.

However, those past studies used existing social categories, with which children were familiar and for which they might have stereotypical expectations (e.g., race, gender, or ethnicity, as in Diesendruck & haLevi, 2006; Diesendruck & Weiss, 2015; Waxman, 2010). It is possible that children readily made similarity inferences about these existing social categories because they had learned that these categories are particularly important in their cultural environment (Baron et al., 2014). It is thus not clear whether children would make similarity inferences about novel social categories – for which they have no prior knowledge – when provided with labels only. Addressing this question is particularly important regarding the

acquisition of stereotypes, as the mere act of labeling social groups might be sufficient for children to generalize properties to group members.

Another way children might learn to associate properties with social categories is through direct observations (Chambers et al., 2008; Gelman & Roberts, 2017; Holland et al., 1986). This theoretical perspective proposes that children generalize properties to category members from statistical evidence, i.e., by observing correlations between category members and certain properties (Bigler & Liben, 2006; A. M. Gonzalez et al., 2017; Riggs, 2019). Similarly, theories of stereotyping suggest that stereotypes are acquired slowly, with accumulated experience over the lifespan (Bigler & Liben, 2006; Devine, 1989; A. M. Gonzalez et al., 2017; Greenwald & Banaji, 1995). Both perspectives draw on models of learning, proposing that children are "sophisticated pattern-detectors" (Gopnik, 2012; Riggs, 2019, p. 67), who first detect regularities in their environment, and subsequently use those regularities to make inferences (Bigler & Liben, 2006; Oakes et al., 2009; Park & Hastie, 1987; Paulus et al., 2011; Sherman et al., 2013; Shutts & Kalish, 2021). In fact, very early on, infants are sensitive to the statistical relationships they observe in their environment. For instance, an important line of research showed that, from a very young age, children are sensitive to observed frequency when predicting an agent's future behavior (e.g., Aloise, 1993; Boseovski & Lee, 2006; Jacobs & Narloch, 2001) and preferences (e.g., Diesendruck et al., 2015; Kushnir et al., 2010; Ma & Xu, 2011; Wellman et al., 2016). In addition, Powell and Spelke (2013) showed that from 7 months, infants expected that group members would behave similarly after observing their common behaviors. Therefore, children's attention to correlations and regularities in their environment (e.g., seeing men working as physicists) might play a crucial role in shaping the content of stereotypes (Bigler & Liben, 2006, 2007 see also Bigler et al., 2001; Brown & Bigler, 2002; Over & McCall, 2018).

Previous research provides evidence that children generalize properties to category members from statistical evidence. For instance, Diesendruck et al. (2015) showed that children are not only able to predict preferences of an individual based on his or her action, but also to generalize a preference to similar others. In their study, 3-4-year-olds were presented with one or two agents who selected five times the same objects from a larger set of objects. The proportion of selected objects relative to the objects available varied across conditions (namely, 18%, 50%, or 100%). Diesendruck et al. (2015) found that children inferred an agent's preference for a kind of objects when he or she selected objects in the nonrandom sampling condition, i.e., the 18% condition (see also Kushnir et al., 2010). Interestingly, when children

saw two similar agents (e.g., two frogs) choosing the same type of objects, children were as likely to generalize the preference to the same agent as they were to generalize to a similar agent (i.e., to another frog) in all three conditions. These findings revealed that observing two similar agents selecting the same five objects led children to generalize the preference to a novel similar agent. In the same line, Riggs (2019) investigated whether children make similarity inferences about social category members based on statistical evidence. In this study, 4- to 8-years-old children were introduced to four members of a novel social category (e.g., a made-up country of origin) and one member from another social category (e.g., another country of origin). Children then observed four people from the same country displaying the same behavior, whereas the fifth individual, coming from another country, engaged in a different behavior. From 4 years of age, children expected that new individuals from the same country would engage in the same behaviors as their fellow category members (see also Riggs & Long, 2020; Roberts, Gelman, et al., 2017). Together, these past studies revealed that, from preschool years, children can rely on their observation of regularities to predict the behaviors and preferences of new category members.

However, these past studies always made verbal reference to the social categories when presenting their regular actions. For example, Diesendruck et al. (2015) emphasized the identity of the new individual, by stating "Look there is another *frog* here!" (p. 374). Similarly, in Riggs' (2019) study, for each category member presented, the country from which the member came was uttered (see also Hoicka et al., 2021; Riggs & Long, 2020; Roberts et al., 2017). This is problematic because verbally marking categories with labels has a privileged effect on how children learn about social categories, implying that similarly labeled individuals are homogeneous (Diesendruck, 2020; Over & McCall, 2018; Shutts & Kalish, 2021). It is thus not clear from past research whether children would learn to associate properties with social categories without verbal reference to the social category. This is a crucial question because it directly addresses the possibility that children might associate properties with social categories by simply observing regularities in their environment (e.g., would children associate girls with childcare after watching TV adverts of girls playing with dolls?)

The above-mentioned theories are also conditioned by some important developmental factors. Prior research assessing the effects of labels on category-based inductions showed that from preschool years, children expect that people identified by the same label would share the same properties (Diesendruck & haLevi, 2006; Diesendruck & Weiss, 2015; Waxman, 2010). These previous studies, however, used existing social categories, with which children were

familiar. Therefore, it is not clear whether preschool- and school-aged children would make similarity inferences based on mere linguistic markers about *novel* social categories, for which they have no prior knowledge. One developmental perspective proposes that, with increasing linguistic competence, children become more sensitive to subtle verbal cues used to refer to social categories, such as category labels (Over & McCall, 2018). Consequently, school-aged children, but not preschoolers, may expect that members of a novel social category share the same properties on the basis of labels only. As for the theoretical perspective proposing that children learn to associate properties with social categories from their observation of regularities, past research revealed that from 4 years old, children can rely on statistical evidence to predict the properties of social category members (e.g., Diesendruck et al., 2015; Riggs, 2019). Moreover, around 4-6 years, children tend to readily infer norms from observing regular behaviors (S. O. Roberts, Gelman, et al., 2017; S. O. Roberts, Ho, et al., 2017). As such, when observing multiple category members doing the same activity, preschoolers may be already inclined to generalize from one category member to another.

The overarching goal of the present study is to deepen our understanding of the conditions under which children start generalizing properties across members of novel social categories. One theoretical perspective holds that noun labels referring to categories signal important and meaningful social divisions, about which there is likely more to learn (Bigler & Liben, 2007; Over & McCall, 2018; S. J. Sherman et al., 2013; Shutts & Kalish, 2021). If minimal linguistic markers are sufficient for children to generalize properties across members of novel social categories, then the way we identify social groups by means of labels might be critical for the acquisition of stereotypes. Other theoretical perspectives also propose that children might attribute properties to categories through direct observations, by detecting regularities in their environment (Bigler & Liben, 2006; Devine, 1989; S. J. Sherman et al., 2013; Shutts & Kalish, 2021). Previous research showed that children rely on statistical evidence to make category-based inferences about verbally marked category members; but would children attribute properties to social categories from mere observation of regularities without labeling? If they do so, then children's environment, and the images they are exposed to, might play a crucial role in the formation of stereotypes. The present study thus assessed these two theoretical perspectives[7] by testing the effects of simple category labels and

[7] Please note that these two theoretical perspectives should not be conceived as mutually exclusive. Rather they describe processes that can equally play a role on children's learning about social categories.

observation of regular actions on preschool- and school-aged children's predictions about novel social categories.

3.2. The present study

To understand the developmental course of similarity inferences in childhood, the present study tested two theoretical perspectives. The first perspective holds that the mere act of labeling social categories is sufficient for school-aged children, but not younger children, to generalize properties to category members (Bigler & Liben, 2007; Over & McCall, 2018; S. J. Sherman et al., 2013; Shutts & Kalish, 2021). The second account proposes that from preschool years children can generalize properties to social category members from their observation of regularities (Bigler & Liben, 2006; Devine, 1989; S. J. Sherman et al., 2013; Shutts & Kalish, 2021). We thus assessed whether preschoolers[8] (4-6 years old) and school-aged children (7-9 years old) would make similarity inferences, i.e., infer that category members share the same properties, based on category labels, observed regularities, or a combination of both. The tasks employed in the present study were inspired by typical category-based induction tasks (Diesendruck & Weiss, 2015; Gelman & Markman, 1986). Children first learned about novel social categories and their properties, and then predicted which of the properties the test exemplars would display. Children also predicted the properties of members from another social category to control that they generalized properties consistently to only members of the same social category, and not to anyone, regardless of category membership. Novel social categories were used because they enable to assess processes that are not tied to a specific cultural background or affected by previous experience with the social category.

In order to assess the first theoretical perspective, i.e., that simple category labels are sufficient for school-aged children to make similarity inferences, we used noun labels (e.g., "a Krollen," "a Trielen") to refer to social category members (see for e.g., Waxman, 2010). Whereas past studies relied on existing social categories, here, we used categories that are not deemed as meaningful and for which children had no prior knowledge to determine whether the mere act of labeling social groups prompts children to make similarity inferences. To investigate the second theoretical perspective, i.e., that observing regularities leads children to make similarity inferences, we presented a majority of unlabeled category members (6 out of

[8] In Germany, preschool years include children between 3 and 6 years old. In the current study, all children of 6 years old had not yet entered elementary school at the time of data collection.

8) sharing the same property. The use of majority (6:8 ratio) rather than totality (8:8 ratio) not only ensured that we were not testing a simple case of associative learning (whereby children would learn to associate properties with anyone, regardless of their category membership), but also provided greater ecological validity (as it is rare to observe absolutely all the members of a category sharing the same properties). In the present study, the way information was presented varied across four conditions that specifically manipulated the two theoretical accounts, namely the use of category labels and direct observation of regularities: (1) Category Label (a category member was introduced with a label), (2) Observed Regularities (eight category members were introduced without label, six of which sharing the same property), (3) Interaction (combination of the Category Label and Observed Regularities conditions: eight category members were introduced with a label, six of which sharing the same property), and (4) Control (one category member was introduced without label).

As specified in our pre-registration (https://osf.io/gnbyp/), we expected that children's prediction scores – for both age groups – would be the lowest in the Control condition, and the highest in the Interaction condition, where both category labels and observed regularities were provided. The Category Label condition tested the hypothesis that the mere act of labeling social category could be sufficient for school-aged children to make similarity inferences, but not for preschoolers (Bigler & Liben, 2007; Over & McCall, 2018; S. J. Sherman et al., 2013; Shutts & Kalish, 2021). We thus expected to find an age group difference in this condition, with 7-9-year-olds having higher prediction scores than 4-6-year-olds. The Observed Regularities condition addressed the second theoretical perspective, postulating that from preschool years children may be already inclined to make similarity inferences from their observation of regular actions (Bigler & Liben, 2006; Devine, 1989; S. J. Sherman et al., 2013; Shutts & Kalish, 2021). Accordingly, we expected that prediction scores of 4-6-year-olds children would be similar to those of 7-9-year-olds in the Observed Regularities condition. In addition, we controlled whether children, depending on their age group and the experimental condition, would make similarity inferences only for members of the same social category, and not over-generalize their predictions to members of other social categories. This latter analysis was exploratory.

3.3. Methods

3.3.1. Participants

A total of 192 German-speaking children (53.6% female), between 4 and 9 years, participated in the study (92 4-6-year-olds and 100 7-9-year-olds). Families were asked to indicate age and gender of their child. Direct question relating to race and ethnicity were not used because these types of questions are very sensitive and unusual in Germany. Due to Covid-19 restrictions, participants were tested online through videoconference and with a Qualtrics (Provo, UT) survey. To take part in the study, families needed an internet connection and a computer/tablet to run the experiments. Only children who spoke fluently German were selected to participate in the experiments. Children were recruited from the city of Munich and its surroundings, on the basis of a lab database of interested families and of a list of families' addresses provided by the city. Families were contacted by e-mail or postal mail. Children's parents gave written consent at the beginning of the study and children orally assented to participate in the study. The study was approved by the institutional ethics committee and complies with American Psychological Association ethical standards. Children received a certificate and a coloring book to thank them for their participation.

Sample size was calculated a priori based on an analysis of variance for a mixed within-between-subjects design (see pre-registration, https://osf.io/gnbyp/). Effect size assumptions were based on a pilot study conducted with 19 children, as well as on other studies that employed similar experimental designs (e.g., Riggs et al., 2014a; Roberts, Ho, et al., 2017; Shilo et al., 2021). The observed standard deviation in these studies varied between .50 and 1.50. The observed standard deviation in our pilot study was similar (1.22). Sample size calculation was thus conducted by setting the standard deviation to 1.00, with an effect size difference of .80, a type I error of .05 and a power of .80. This analysis led to a sample size estimation of 192 children. No additional participant was recruited once the predetermined sample size was reached.

As specified in the pre-registration, we excluded data from participants who stopped being attentive to the task or failed to complete the study (n = 6), whose responses were influenced by parents or siblings (n = 2), or who did not provide a correct answer to the

"belonging check" question (n = 4). The final sample size resulted in 180 children[9]. Children were divided into two age groups: 4- to 6-year-olds (n = 88, 53.4% female, age mean = 5.49, range = 4 years 0 month to 6 years 11 months) and 7- to 9-year-olds (n = 92, 56.5% female, age mean = 8.50, range = 7 years 0 month to 9 years 11 months). The Category Label condition included 59 children (4-6 years: n = 29, age mean = 5.42; 7-9 years: n = 30, age mean = 8.42), the Observed Regularities condition included 59 children (4-6 years: n = 29, age mean = 5.46; 7-9 years: n = 30, age mean = 8.53), and the Interaction condition included 62 children (4-6 years: n = 30, age mean = 5.57; 7-9 years: n = 32, age mean = 8.55).

3.3.2. Material and procedure

3.3.2.1. Stimuli

The stimuli consisted of short cartoon movies and were created with Microsoft PowerPoint. The cartoon illustrations were modified versions of the stimuli used by Moty & Rhodes (2021). A pre-recorded narrator guided children through the different tasks. In each task, children learned about novel categories of people distinguished by the country they came from, the language they spoke, and the types of hats they wore. Characters within each category were different with respect to their gender, ethnicity, morphology, and clothing style (see Figure 1). All study materials are available online at https://osf.io/567rw/.

3.3.2.2. Procedure

Upon starting the experiments, families were provided with guidelines on how to set the computer/tablet and instructions regarding parents' role during the study (e.g., parents could intervene in case of technical issues or to encourage their child, but not interfere with their child's answers). An experimenter was present through videoconference during the entire duration of the study in order to control for parents/siblings' interferences, children's attention during the tasks, and to provide support when needed. Before starting the tasks, the experimenter made sure that parents and children had understood the procedure and answered any additional questions. The experiment lasted approximately 15 minutes.

Warm-up task. Children first completed a warm-up task to familiarize themselves with the procedure. The warm-up task started with a female narrator, introducing children to a special

[9] A post hoc analysis of power with the current sample size revealed that a power of .77 was achieved with the final sample of 180 subjects.

and distant town, where different categories of people live. Children were informed that they would learn about the activities of the people living in this town, and were directly presented with an example (e.g., "*Diese Person pflückt diese Blume*."; "This person picks this flower"). Then a new character from the town was displayed, and children were asked to choose which type of flower, out of three choices, this new individual would pick (e.g., "*Was denkst du, welche Blume pflückt diese neue Person?*"; "What do you think, which flower does this new person pick?"). All children passed the warm-up task.

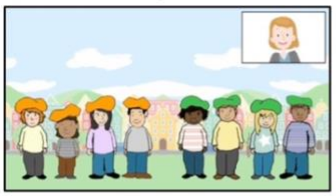
Introduction Phase. Children learned about two categories of people that were distinct with respect to the country they came from, the language they spoke, and the types of hats they wore (see Figure 1 for a summary of the protocol). For the Category Label and Interaction conditions, a category label was uttered twice to differentiate category members (e.g., "*Die Krollen kommen aus einem Land mit vielen Wüsten. Die Krollen tragen grüne Hüte und sprechen diese Sprache. Hör zu!*"; "The Krollens come from a country with many deserts. The Krollens wear green hats and speak this language, listen!"). Children then learned about the characteristics (i.e., language, country of origin, and hat color) associated with the other social category. Hats' colors (red, blue, green, or orange), spoken languages (speech extracts of Chinese, Russian, Antillean Creole from Guadeloupe or Italian), the country's specificities (with a lot of beaches, forests, deserts, or mountains), the position on screen (left or right), and if applicable, the label ("Trielens" or "Krollens") associated with each social category were fully counterbalanced. Hat colors, and if applicable category labels, were used during each phase of the task to mark the two social categories. Country of origin and language were used only in the "Introduction Phase" to ensure that children considered the two sets of people as distinct social categories. This decision was motivated by the fact that very early on, children can rely on language and accent to form and distinguish social categories (Kinzler et al., 2007, 2010; Liberman et al., 2017), but it is only from 5 years of age that children start to consider clothing as marking social category boundaries (Baron et al., 2014; Diesendruck & Weiss, 2015). The speech extract for each selected languages consisted in recordings of native speakers[10]. To further control that children correctly learned to distinguish the two social

[10] The speech extracts were modified (chunks of recording were cut and pasted at different places in the track). The goal of this modification was to avoid any possible risk that if some participants spoke one of the languages played during the experiment, they would immediately assume that the group presented in the experiment was the real existing linguistic category. In addition, we also asked parents to indicate with which languages their child interacted in their daily life to further ensure that the categories presented during the experiment remained novel for the child. No participant was reported to have daily interactions with any of the languages selected for the present study.

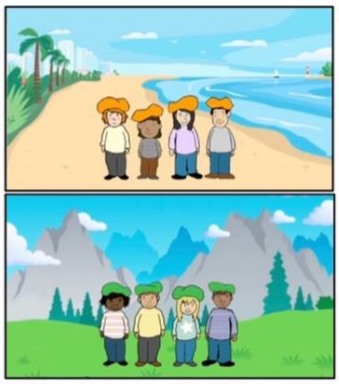
categories, children in the present study answered two "belonging check" questions. Specifically, children were asked to link two new individuals to their corresponding category (see Figure 1, vignette 3). Children providing an incorrect answer were excluded from the study ($n = 4$).

Figure 1. Protocol for the Introduction, Learning, and Test phases.

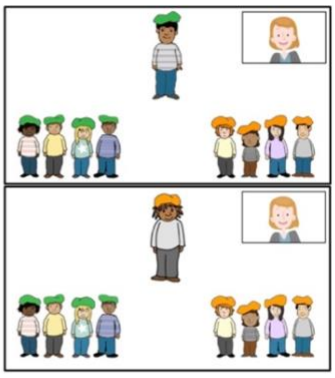
Introduction phase



(1) The narrator introduced the people who live in the special and distant town.

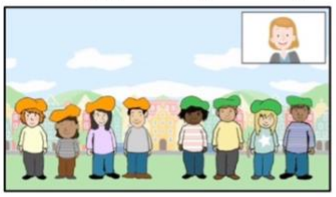


(2) The narrator highlighted the distinctions between the two novel social categories (hat colors, country, language)
- Control / Observed Regularities: no label ("these people")
- Label / Interaction: category labels ("Krollens"; "Trielens")

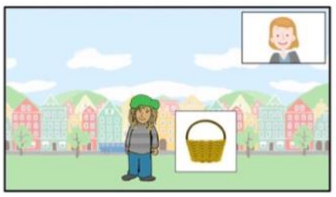


(3) Belonging check question: Participants were asked to link the new individuals to their corresponding social category.

Learning phase (4 trials)



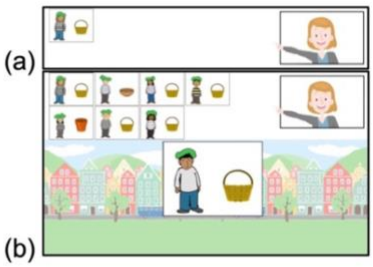
(4) The narrator drew attention to the social category about which participants would learn (here, the green hats, on the right)



(5) The narrator introduced a new category member and her activity

- Control / Observed Regularities: no label ("there is a new person")

- Label / Interaction: with label ("there is a new Krollen")

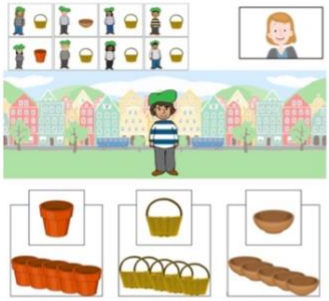


(6) The narrator created a memory cue to remind participants of the property of each category member.

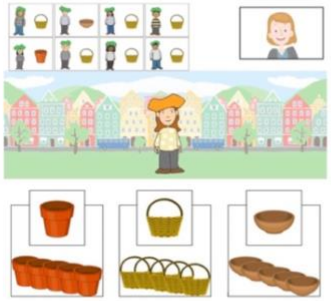
(a) Control and Label conditions

(b) Observed Regularities and Interaction conditions

Test phase (4 trials)



(7a) Prediction for the *same* category: participants attributed a property to a member of the same social category as presented in the learning phase.
- Control / Observed Regularities: no label ("there is a new person")
- Label / Interaction: with label ("there is a new Krollen")



(7b) Prediction for the *other* category: participants attributed a property to a member of the other social category.
- Control / Observed Regularities: no label ("there is a new person")
- Label / Interaction: with label ("there is a new Trielen")

Learning Phase. Children learned about the activities performed by the members of one of the two social categories (see Figure 1, vignettes 5 and 6). There were four trials, each of which presenting one activity type. The four activity types were games (puzzles, dominos, or board games), food (peas, broccolis, or green beans), transport (scooter, roller-skates, or skateboards), and manufactured objects (pots, bowls, or baskets). The selected properties were activities because they are neither considered as too subjective (such as preferences) nor as too generalizable (such as morphological properties) (Diesendruck & Eldror, 2011; Doan et al., 2021; Riggs et al., 2014a). The way information was presented during the "Learning Phase" varied across conditions (see below). All children first completed the Control Condition before being randomly assigned to one of the test conditions, i.e., Category Label, Observed Regularities, or Interaction Condition. In order to ensure that children's inferences were not influenced by repetition effects, the activities and the social categories systematically differed across the Control Condition and the Test Condition. That is, children made inferences about two different social categories and never learned about the same activity between these conditions. For example, if children learned in the Control Condition that the activity of the first social category (e.g., the red hats) was making baskets, then the activity of the second social category (e.g., the green hats) in the Test Condition was necessarily either making pots or bowls, but never baskets. In this way, information was never repeated between the Control and the Test conditions.

Control Condition. At the beginning of each trial, the narrator informed children about which characters they were going to learn (e.g., "*OK, jetzt werden wir erfahren welche Beschäftigungen diese Personen haben.*"; "Okay, now we are going to learn about the activities of these people [*animation was played on the green hats group*]"). The two social categories disappeared, and a new category member, visually marked by the color of his or her hat (e.g., green hat), appeared on the screen. Children then learned about the activity of this new person (e.g., "*Schau mal, das ist eine neue Person. Jetzt werden wir erfahren was sie/er herstellt.*"; "Look, there is a new person [*the category member appears on screen*]. Now we'll find out what she/he makes [*the activity's picture, e.g., a basket, pops up next to the category member*]"). The category member and the picture of his or her activity then moved to the top left corner of the screen to serve as a memory cue during the "Test Phase." The narrator explicitly informed children about this memory cue (e.g., "*Ich verschiebe, das Bild nach oben um dich daran zu erinnern, was sie/er tut.*"; "I move the picture on top to remind you what she/he does") before moving to the "Test Phase."

Category Label Condition. This condition is the same as the Control Condition except that the categories were introduced with a label (e.g., "*Schau mal, das ist ein neuer Kroller. Jetzt werden wir erfahren was er herstellt*."; "Look, there is a new Krollen [*the member, e.g., with a green hat, appears on screen*]. Now we'll find out what he makes [*the activity's picture, e.g., a basket, pops up next to the Krollen*]"). This condition assessed the theoretical perspective that the mere act of labeling social categories leads children to generalize properties to category members (Bigler & Liben, 2007; Over & McCall, 2018; S. J. Sherman et al., 2013; Shutts & Kalish, 2021). Again, a memory cue was provided before moving to the "Test Phase."

Observed Regularities Condition. Children learned about the activities of eight category members, one at a time (e.g., "*Schau mal, das ist eine neue Person. Jetzt werden wir erfahren was sie/er herstellt.*"; "Look, there is a new person [*the category member, e.g., with a green hat, appears on screen*]. Now we'll find out what she/he makes [*the activity's picture pops up, e.g., a basket*]" repeated for the eight members). To address the research question that children can generalize properties to category members from their observation of regularities (Bigler & Liben, 2006; Devine, 1989; S. J. Sherman et al., 2013; Shutts & Kalish, 2021), we presented six category members with the same activity (e.g., six members made baskets) and two members having different activities (e.g., one person made pots, and one person made bowls). The presentation order of each activity was randomized. A memory cue was provided for each category member, resulting in 8 memory cues displayed on the top left of the screen (see Figure 1, vignette 6b).

Interaction Condition. This condition is the same as the Observed Regularities Condition except that the categories and each member were introduced with a label (e.g., "*Schau mal, das ist ein neuer Kroller. Jetzt werden wir erfahren was er herstellt*."; "Look, there is a new Krollen [*the member appears on screen*]. Now we'll find out what he makes [*the activity's picture pops up*]" repeated for eight category members). In this way, we assessed whether verbal reference to the social category is necessary for children to make similarity inference.

Test Phase. Immediately after each trial of the "Learning Phase," the narrator invited children to predict the activities of a member from the *same* social category, and the activity of a member from the *other* social category (see Figure 1, vignettes 7a and 7b). In the Category Label and Interaction Conditions, category membership was marked by hat colors and labels (e.g., "*Schau mal ein neuer Kroller/Trieler. Was denkst du, was stellt er her?*"; "Look, a new

Krollen/Trielen! What do you think, what does he make?"), whereas in the Control and Observed Regularities Conditions, category membership was only visually marked by hat colors (e.g., "*Schau mal eine neue Person. Was denkst du, was stellt er her?*"; "Look, a new person! What do you think, what does he make?"). Thus, children made predictions about an individual belonging to the *same* social category as the one presented in the "Learning Phase," as well as about an individual belonging to the *other* category (presented in the "Introduction Phase"). The order of the questions was counter-balanced. The predictions about the *same* category members enabled us to determine if children inferred that individuals would engage in the same activities based on category membership. The predictions about the *other* category members enabled us to assess if children over-generalize the activity to members of another category, and not consistently to only members of the same social category. Children could select only one of the 3 possible answers for each type of activity (e.g., "pots," "bowls," or "baskets" for the type of made objects).

3.3.3. Measures and coding

Each condition consisted in four trials, corresponding to the predictions about the four activities (games, food, transport, and object making). Children were assigned to the Control Condition and then, randomly to one of the test conditions, resulting in a total of 8 trials. For each trial, children made two category-based inductions: one about a member of the *same* category, and one about a member of the *other* category.

A prediction score was computed for each participant. Predictions were scored as 1 for a consistent response, and as 0 for an inconsistent response. A consistent response consisted in predicting the same property as the one presented in the "Learning Phase." For example, in the Interaction Condition, participants could observe six out of eight Krollens making baskets during the Learning Phase. A consistent response for the *same* category members, in this case, consisted in predicting that a new Krollen also made baskets. If participants predicted that a new Krollen made pots or bowls, they received no point. Similarly, a consistent response for the *other* category members, in the above example, was to predict that a Trielen made baskets; no point was given if participants predicted that the Trielen made pots or bowls. Overall, participants could obtain a maximum score of 4 points per condition and per prediction type (i.e., member from the *same* or the *other* category).

3.3.4. Transparency and openness

The study design, sampling and analysis plans were pre-registered before data collection at https://osf.io/gnbyp/. We report how we determined the sample size, data exclusions, all manipulations, and all measures in the study. The materials, data and analyses' script have been made publicly available and can be accessed at https://osf.io/567rw/. The data analyses and plots were performed with RStudio, version 2022.2.2.485 (RStudio Team), using the additional packages *afex* (Singmann et al., 2016) and *ggplot2* (Wickham, 2016).

3.4. Results

3.4.1. Difference between the control condition and the test conditions

The Control Condition represents baseline prediction scores, as children could have used only visual cues (hat colors) to predict the properties of new category members. Comparing prediction scores for the *same* category member in the Control Condition to absolute chance (1/3 selection of the consistent property across four trials, resulting in chance = 1.333) revealed that participants in the Control Condition made category-based inferences significantly less than expected by chance (M = 1.11, SD = 1.15), $t(179) = -2.66$, $p = .009$, $d = -0.20$. To analyze if there is a difference between the Control Condition and the test conditions, i.e., that the experimental manipulation (Category Label, Observed Regularities, and Interaction Conditions) had an effect on participants' similarity inferences about members of the *same* category, we conducted three dependent *t*-tests on prediction scores, applying a Holm-Bonferroni correction (see pre-registration, https://osf.io/gnbyp/). Predictions scores were significantly higher in the test conditions compared to the Control Condition (see Table 1).

Table 2. Prediction scores for the same social category member in the control condition compared to each test condition.

Condition	*M* (*SD*) control	*M* (*SD*) test	df	*t*	*p*	*d*
Control - Category Label	1.08 (1.29)	1.63 (1.35)	58	4.37	<.001*	.57
Control - Observed Regularities	1.14 (1.17)	1.71 (1.29)	58	3.30	.002*	.31
Control - Interaction	1.10 (1.00)	2.00 (1.37)	61	4.46	<.001*	.36

Note. Statistics paired *t*-tests comparing each test condition with the control condition.
* *p*-value < adjusted *p*-value with Holm-Bonferroni correction (rank 1: adjusted *p*-value = .017; rank 2: adjusted *p*-value = .025; rank 3: adjusted *p*-value = .05).

3.4.2. Predictions for the same social category member

We analyzed whether age groups and test conditions would impact children's category-based inductions. That is, we assessed whether children would attribute more often the same activities to members of the same social category depending on their age group and the test condition. As planned in the pre-registration (https://osf.io/gnbyp/), we performed a 2 (age groups: 4-6, 7-9 years old) x 3 (test conditions: category label, observed regularities, interaction) factorial ANOVA with prediction scores for the *same* category member as dependent variable. The factorial ANOVA revealed a main effect of age groups, $F(1, 174) = 6.78$, $MSE = 1.72$, $p = .010$, $\eta_p^2 = .04$, with 7-9-year-olds having higher prediction scores ($M = 2.03$, $SD = 1.43$) than 4-6-year-olds ($M = 1.52$, $SD = 1.18$). There was however no main effect of the test conditions (category label, observed regularities, and interaction), $F(2, 174) = 1.38$, $MSE = 1.72$, $p = .270$, $\eta_p^2 = 0.02$, or interaction effect between age group and test conditions, $F(2, 174) = 1.28$, $MSE = 1.72$, $p = .280$, $\eta_p^2 = 0.02$.

Although no interaction effect between age and condition was observed, we nonetheless checked age group differences for each test condition, separately, as predefined in our pre-registration (even if the separate analyses were not strictly justified by the ANOVA). We thus ran independent *t*-tests for each condition, comparing prediction scores of 4-6-year-olds with those of 7-9-year-olds (see Figure 2). Because 3 independent *t*-tests were run, we again applied a Holm-Bonferroni correction for multiple hypotheses testing (see pre-registration, https://osf.io/gnbyp/). The resulting adjusted *p*-values are .017, .025 and .050 for each rank respectively.

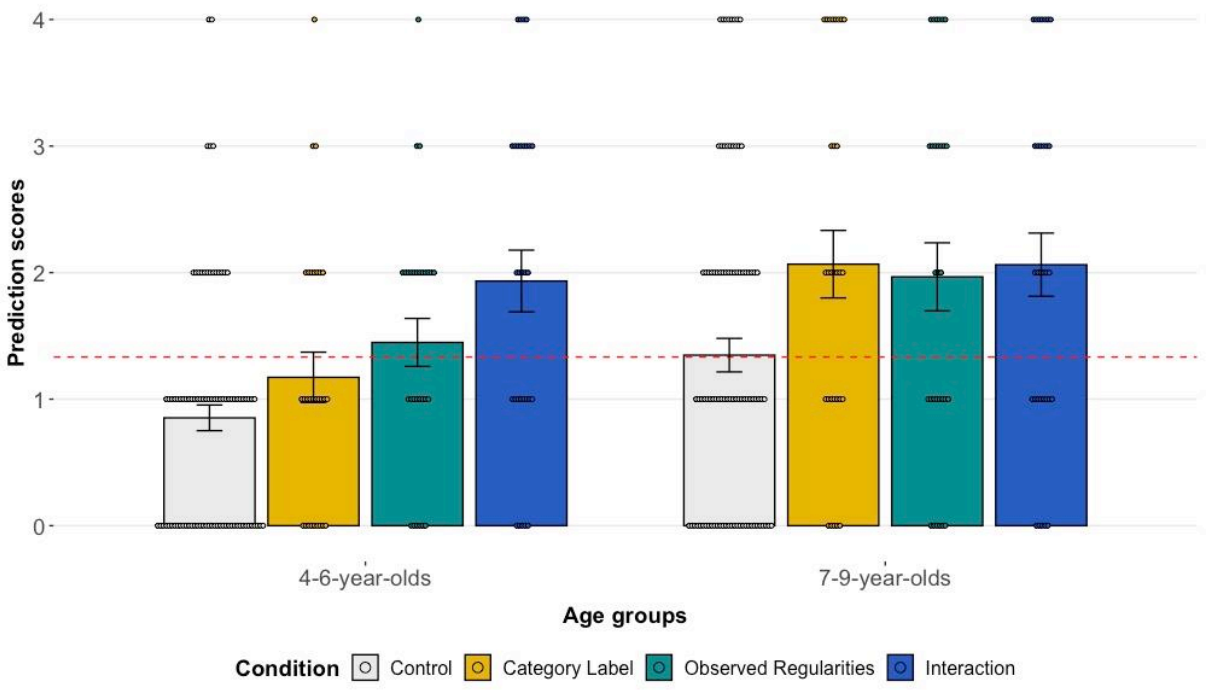
For the Category Label condition, we expected that prediction scores of 7-9-year-olds would be higher than prediction scores of 4-6-year-olds. The *t*-test confirmed this hypothesis: 7-9-year-olds achieved higher prediction scores ($M = 2.07$, $SD = 1.46$) than 4-6-year-olds ($M = 1.17$, $SD = 1.07$), $t(57) = -2.67$, $p = .010$ (< adjusted $p = .017$), $d = 0.70$. When comparing younger and older children's predictions against chance, we found that younger children's performance was not different from chance level ($t(28) = -0.81$, $p = .425$, $d = -0.15$), whereas older children performed significantly above chance ($t(29) = 2.75$, $p = .010$, $d = 0.50$) (see Figure 2).

For the Observed Regularities condition, we expected that prediction scores of 4-6-year-olds would be similar to prediction scores of 7-9-year-olds. The independent *t*-test revealed no significant difference between the prediction scores of the 4-6-year-olds ($M = 1.45$, $SD = 1.02$)

and the 7-9-year-olds (M = 1.97, SD = 1.47), $t(57)$ = -1.57, p = .123 (> adjusted p = .025), d = 0.41. Younger children's predictions were not different from chance level ($t(28)$ = 0.61, p = .549, d = 0.11), but prediction scores of older children were significantly above chance ($t(29)$ = 2.35, p = .026, d = 0.43) (see Figure 2).

For the Interaction condition, we expected that prediction scores would be similar between the two age groups. Prediction scores of 4-6-year-olds (M = 1.93, SD = 1.34) and of 7-9-year-olds (M = 2.06, SD = 1.41) were in fact similar to each other, yielding no significant difference, $t(60)$ = -0.37, p = .713 (> adjusted p = .05), d = 0.09. Predictions scores of both age groups were significantly above chance (younger group: $t(29)$ = 2.46, p = .020, d = 0.45; older group: $t(31)$ = 2.92, p = .006, d = 0.52) (see Figure 2).

Figure 2. Prediction scores per conditions and age groups for members of the same social category.



Note. Mean prediction scores of 4-6-year-olds and 7-9-year-olds in the Control (very light gray), Category Label (yellow/light gray), Observed Regularities (green/gray), and Interaction (blue/dark gray) conditions, with datapoints displayed. Error bars represent standard errors of the mean. The red dashed line corresponds to absolute chance level (i.e., prediction scores of 1.33).

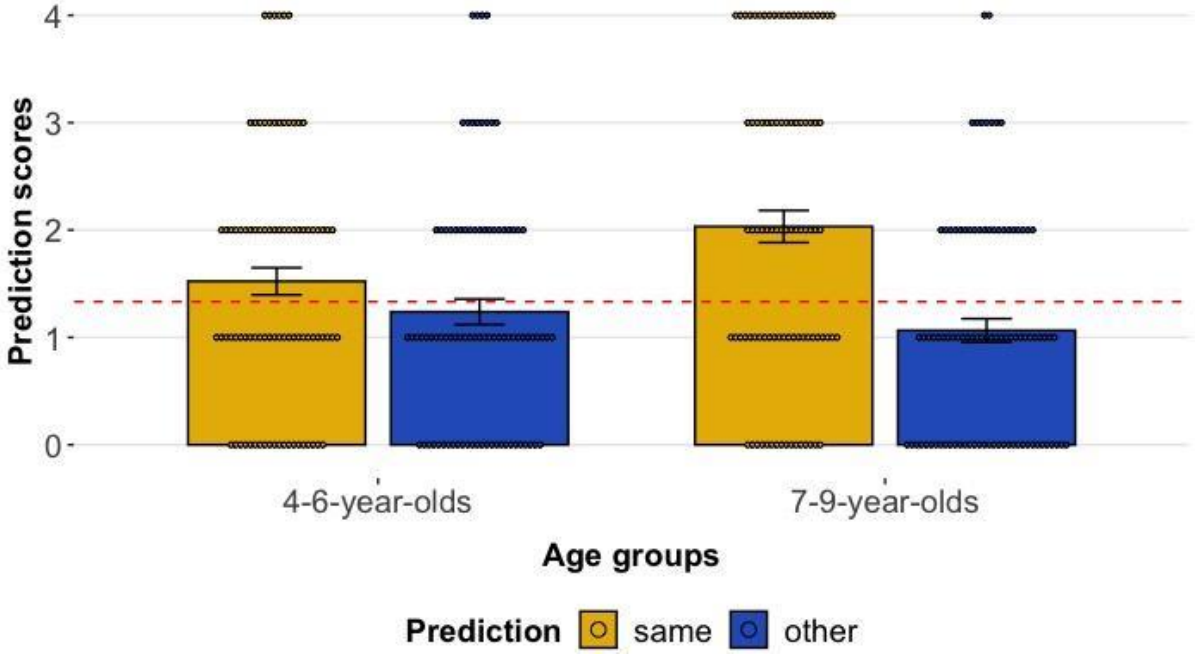
3.4.3. Do children over-generalize their predictions to members of another social category?

To explore whether children over-generalized their predictions to members of another category, we performed a mixed 2 (age groups: 4-6 years, 7-9 years) x 3 (test conditions: category label, observed regularities, interaction) x 2 (category: same, other) ANOVA with prediction scores as dependent variable. The ANOVA revealed a main effect of category, $F(1, 174) = 28.29$, $MSE = 1.24$, $p < .001$, $\eta_p^2 = 0.14$, with prediction scores that were higher for the *same* ($M = 1.78$, $SD = 1.34$) than for the *other* category member ($M = 1.15$, $SD = 1.08$). There was an interaction effect between age group and category, $F(1, 174) = 8.54$, $MSE = 1.24$, $p = .004$, $\eta_p^2 = 0.05$, which was the only other effect to achieve significance (all other ps > .127).

We conducted post-hoc analyses to further explore the difference between younger and older children's predictions for the *same* and *other* category members. Because there was no effect of condition, we collapsed prediction scores across test conditions. Post-hoc analyses revealed that 7-9-year-olds predicted consistently that only members of the same social category would share the same properties. Their prediction scores for the *same* social category member were significantly higher than for the *other* social category member, $t(91) = 5.44$, $p < .001$, $d = 0.57$. Moreover, when comparing 7-9-year-olds' prediction scores against chance, we found that 7-9-year-olds made significantly less similarity inferences for the *other* category member than expected by chance ($M = 1.07$, $SD = 1.05$), $t(91) = -2.46$, $p = .016$, $d = -0.26$. For the *same* category member, however, 7-9-year-olds made category-based inferences significantly more than expected by chance ($M = 2.03$, $SD = 1.43$), $t(91) = 4.68$, $p < .001$, $d = 0.49$ (see Figure 3).

For 4-6-year-olds, there was no significant difference between their predictions for the *same* and the *other* social category member, $t(87) = 1.91$, $p = .060$, $d = 0.20$. Their predictions for the member of the *same* ($M = 1.52$, $SD = 1.18$) and *other* ($M = 1.24$, $SD = 1.11$) social category were not significantly different from chance level (same: $t(87) = 1.50$, $p = .137$, $d = 0.16$; other: $t(87) = -0.80$, $p = .428$, $d = -0.08$) (see Figure 3).

Figure 3. Prediction scores for the same and the other category members per age groups.



Note. Mean prediction scores of 4-6-year-olds and 7-9-year-olds obtained for the *same* (yellow/light gray) and the *other* (blue/dark gray) category members, with datapoints displayed. Error bars represent standard errors of the mean. The red dashed line corresponds to absolute chance level (i.e., prediction scores of 1.33).

3.5. Discussion

The current research investigated two theoretical perspectives proposing that children learn to associate properties with social categories from *simple category labels* (Bigler & Liben, 2007; Over & McCall, 2018; S. J. Sherman et al., 2013; Shutts & Kalish, 2021), or from their *observation of regularities* (Bigler & Liben, 2006; Devine, 1989; S. J. Sherman et al., 2013; Shutts & Kalish, 2021). It took a developmental perspective by comparing preschool- (4-6 years) and school-aged (7-9 years) children's similarity inferences about social category members. More concretely, we assessed whether children predicted that an individual would engage in the same activity as its fellow-category members after being exposed to (a) simple linguistic markers of social category, i.e., one category member referred to by a label, (b) mere observation of regularities, i.e., observing a majority of unlabeled category members doing the same activity, (c) simple category labels and observed regularities, i.e., observing a majority of labeled category members doing the same activity, and (d) a control condition, i.e., a single category member introduced without label.

Overall, the present study supports and extends the theoretical perspective that children learn to associate properties with social categories from their observation of regularities (Bigler & Liben, 2006; Devine, 1989; S. J. Sherman et al., 2013; Shutts & Kalish, 2021). Previous research revealed that from 4 years old, children can rely on statistical evidence to predict the properties of social category members (e.g., Diesendruck et al., 2015; Riggs, 2019). However, these past studies marked verbally the identity of category members upon presenting their regular actions. In the present study, two conditions were created to assess separately whether children would associate properties with social categories from their observation of regular actions, with or without category labels. On the one hand, 7-9-year-olds, but not 4-6-year-olds, made similarity inference from observing a majority of *unlabeled* category members engaging in the same activity. On the other hand, and similar to previous research (Diesendruck et al., 2015; Riggs, 2019; S. O. Roberts, Gelman, et al., 2017), from preschool age, children made similarity inferences from statistical evidence when labels were provided. These findings support the theoretical perspective that, as early as the preschool years, children generalize properties to verbally marked category members from their observation of regularities (Bigler & Liben, 2006; Devine, 1989; S. J. Sherman et al., 2013; Shutts & Kalish, 2021). Moreover, the present study adds novel perspective to this theoretical account, showing that around 7-9 years, children start to make similarity inferences from the regularities they observe between properties and category members, without needing labels.

The present study also confirms the theoretical perspective that the mere act of labeling social categories is sufficient for school-aged children, but not preschoolers, to generalize properties to category members (Bigler & Liben, 2007; Over & McCall, 2018; S. J. Sherman et al., 2013; Shutts & Kalish, 2021). From 7-9 years, children could make similarity inferences based on simple category labels; younger children failed to generalize properties from simply learning about the property of *one* labeled individual. These results complement previous findings that showed that preschoolers are prone to generalize properties across category members (e.g., Diesendruck & haLevi, 2006; Diesendruck & Weiss, 2015; Kalish, 2012; Rhodes & Gelman, 2008; Waxman, 2010), and should be interpreted in the light of two important considerations. First, the present study differs from previous research in the degree to which the tested social categories were essentialized. Previous research revealed that from preschool years, children readily associate attributes with *existing* social categories (e.g., Diesendruck & haLevi, 2006; Gelman et al., 1986; Rhodes & Gelman, 2008; Waxman, 2010, for a review see Charlesworth & Banaji, in press). This is because existing social categories

tend to be highly essentialized, leading children to assume that category members share the same properties because of an underlying essence (Diesendruck, 2020; Mari, 2022). Children, however, do not hold these kinds of essentialist beliefs about novel artificial categories (Hoicka et al., 2021; Leshin et al., 2021; Rhodes, 2013; Vasilyeva et al., 2018). Second, the present study assessed whether children could draw similarity inferences from mere labels, rather than labels in combination with generic language. As previously discussed, generic language constitutes a privileged cultural input that guides children's inferences and essentialist beliefs about social categories (Bloom, 2004; Cimpian & Markman, 2011; Gelman, 2003). And indeed, a large number of studies documented the fact that using labels with generic statements prompted preschoolers to assume that *any* properties are generalizable to members of the same *novel* social category (e.g., Baron et al., 2014; Jordan & Dunham, 2021; Moty & Rhodes, 2021; Rhodes et al., 2012, for a review, see Mari, 2022). The present study thus specifies previous findings, showing that the mere act of labeling individuals might not be sufficient for young children to attribute properties to category members of novel artificial social categories. Yet, as children develop increased linguistic competence and experience school settings and peer interactions, they also become more sensitive to mere verbal cues to social categories (Over & McCall, 2018). In other words, from early school years, children start to consider mere labels as meaningful markers of social divisions.

In addition, our study provides some clarifications regarding the extent to which children make similarity inferences. To our knowledge, previous studies did not assess whether children generalize properties to only members of the *same* social category or if they tend to *over*-generalize the properties to any individuals, regardless of their social category. This distinction is theoretically important in determining whether the social category *per se* had inductive force, and in ruling out the possibility that children projected properties onto people in general (Shutts et al., 2013; Waxman, 2010). An exploratory analysis was conducted to assess this question and revealed that 7-9-year-olds predicted that only members of the same social category would engage in the same activities; their predictions for the members of another social category were at chance level. Younger children, on the other hand, failed to expect that only members of the same social category would engage in the same activities. Overall, these results suggest that from 7-9 years old, children start to generalize properties consistently to only members of the same novel social categories from minimal linguistic markers or mere observation of regularities.

One could argue that the seemingly low generalization rates observed among younger children might be due to the online administration of the tasks, which might have had an impact on younger children's attention. However, we doubt that this was the case as several strategies have been applied to ensure that children would remain attentive during the task and to reduce the possible demands of the study. First, we provided breaks between trials, thereby offering children the possibility to make a pause if they needed to. Second, for the younger children, parents were present to help them using the computer/tablet and, if necessary, to support and encourage them during the task. Third, the experimenter monitored that children were attentive during the whole duration of the study. Data of children who stopped being attentive to the task were excluded (cf. Methods' section). What is more, a growing number of research in developmental psychology started to evaluate the validity of online testing, and demonstrated that online studies yielded similar results as in-person studies (for e.g., Rhodes et al., 2020; Schidelko et al., 2021; Vales et al., 2021; Yamamoto et al., 2021). In particular, Rhodes et al. (2020) conducted online several studies that were originally face-to-face, with similar paradigms as the one presented here, and with children of the same age range (between 4 and 8 years old). Their online replications yielded similar results as the in-person studies, suggesting that these kinds of paradigms (i.e., where children listen to simple stories and respond by choosing between pictured response options) are well suited for online testing with children from 4 years old (Rhodes et al., 2020, see also Schidelko et al., 2021).

In future work, it will be important to examine the interaction between different type of properties and category-based induction processes. In fact, previous research showed that certain properties are more likely generalized across category members than others. For example, children are prone to expect that members of novel social categories share morphological features and psychological properties (Diesendruck & Eldror, 2011; Riggs et al., 2014a) but not preferences or historical properties (Doan et al., 2021; Riggs et al., 2014a). In the present study, we tried to use properties that offered a trade-off between very subjective properties (e.g., preferences) and very generalizable properties (e.g., morphological features, psychological traits). It is possible that preschoolers might be more inclined to make similarity inferences with properties that are considered as more generalizable. For example, using less common, more distinctive properties, such as unusual behaviors (e.g., "raising arms," Riggs, 2019) or novel properties (e.g., "eating a new snack called *naggle*," Waxman, 2010), might lead younger children to make more similarity inferences. In fact, determining which properties are better source of inductive potential represents an interesting avenue for future research (cf.

Diesendruck & Eldror, 2011). In the future, research should take into account the large variety of attributes that can be associated with social category members (e.g., traits, roles, skills, behaviors, beliefs, preferences, interpersonal relations, or physical features), and systematically assess how the type of attributes modulates category-based inferences.

Furthermore, one question raised by our experimental design is whether the choice to use familiar properties without additional element of contrast evidence played a role on the overall pattern of the results. Because the properties used in the present study were familiar, children might have prior biases regarding these properties, inferring that people would display each property to similar extent. As such, receiving evidence that category members engage in one activity (e.g., "the Krollens make baskets"), without contrasting evidence (e.g., "the Krollens *do not* make pots or bowls") might have confirmed children's preconceptions and might have been insufficient to compete with their prior expectations, thereby leading to limited generalization rates. In fact, previous studies, that presented positive *and* negative evidence about the properties of social categories, found overall higher generalization rates among children between 4 and 6 years old (e.g., Foster-Hanson et al., 2021; Kalish, 2012; Roberts, Gelman, et al., 2017; Roberts, Ho, et al., 2017). In these studies, children simultaneously learned about the contrasting properties of two different social categories. This setup might have prompted children to perceive some degree of competition between the social categories, which would in turn increase their generalization of properties to members of a social category (Rhodes & Brickman, 2011; Riggs & Long, 2020). The current research, however, built on the theoretical perspectives that children can learn about the properties of a social category from mere labels or from their observation of regularities. The exact role of contrasting negative evidence in this learning process remains unclear. It is possible that the youngest children would make more similarity inferences based on small evidence, if they were provided with additional information that explicitly presents the properties as mutually exclusive (see for e.g., research on general word learning, Markman, 1987; Markman & Wachtel, 1988; Zosh et al., 2013). Future research should thus systematically consider the role of prior biases and the need of contrasting negative evidence when assessing children's inductive inferences about social categories.

Two limitations to the current research should also be mentioned. First, the present study lacks a better appreciation of the amount of statistical evidence at which children start to make similarity inferences. Because the study was conducted online, we had to limit the duration of the experiment to keep children motivated, and therefore the number of observations had to be

restricted to eight category members. We have to leave it to future research to test varying frequencies, such as 60%, 70%, 80%, or 90%, and increase the number of observed instances (e.g., presenting 10, 15, 20 category members) to determine if there is a threshold from which children start to make similarity inferences based on their observation of regularities. Second, the data derived from a sample of European children, and therefore prevent us from drawing conclusions about children's similarity inferences across human populations. Future cross-cultural research, including underrepresented populations, would provide a better and more nuanced understanding of how children learn about social groups.

In sum, the capacity to make similarity inferences about social categories has cognitive advantages, reducing the complexity of the social world and allowing children to predict and explain category members' features and behaviors (Baron, 2009; Rhodes, 2013; Rhodes & Baron, 2019), but it also sets the ground for the development of stereotypes (Bigler & Liben, 2006; Rhodes & Baron, 2019; Tajfel, 1970). The present study suggests that early school years might be a critical period in which children start to easily attribute properties to social categories from minimal cues. This finding bears some important implications. First, the mere act of labeling social groups (e.g., "Girls go to the right, boys to the left", "Joe is a vegetarian") might prompt school-aged children to infer that similarly labeled people will share the same properties. Second, school-aged children seem also prone to attribute properties to social categories by observing regularities in their environment (e.g., observing a majority of men physicists, of shy librarians, etc.). Third, young children were more conservative in their generalization, attributing properties to social category members on the basis of labels *and* regular actions. This reluctance to make similarity inferences from small evidence could in fact prevent young children from making overly broad generalizations in the process of early conceptual development (cf. Brandone et al., 2015). As such, it seems particularly important to monitor how social categories are represented and referred to during these age periods to possibly reduce some stereotyping effects.

Taken together, the present study's results suggest that from 7-9 years, children readily pick up information about social categories, and infer that new individuals, belonging to the same category, will be similar. The present study suggests that this tendency to rapidly make similarity inferences develops during the early school years and may constitute the ground from which children acquire stereotypes.

Summary and comments

Chapter 2 presented an experimental study that assessed whether children would learn to associate properties with social categories from simple labels or from their observation of behavioral regularities. From around 7-9 years, children generalized activities to category members based on simple labels or on their observations. Four- to 6-year-old children, however, made similarity inferences only when both labels and statistical evidence were combined. While this study provides interesting insights into how children learn to associate properties with social category members, some important questions remain open.

First, a large variety of attributes can be associated with social categories (e.g., traits, skills, beliefs, behaviors, preferences, interpersonal relations, or physical features). Future research should seek to deter-mine which attributes are better source of inductive potential and whether the type of attributes might modulate category-based inferences.

Second, shared properties can be used as a basis for categorization and as a basis for inductive inference. As such, the role of categorization in promoting similarity inferences is not clear: Do children need to rely on category membership to make similarity inferences or could they simply draw on people's similarities to predict that they share other attributes?

Finally, one might question whether the capacity to infer similarity from small evidence is specific to the social domain or whether it also applies to object and animal categories. These remaining questions represent great opportunities to better apprehend how children make inferences about social categories and whether this capacity is specific to the social domain.
The following chapters turn to stereotypes held by adults: Chapter 3 introduces the concept of stereotype and delineates the theoretical framework adopted in the experimental studies presented in Chapter 4 and Chapter 5.

4. Chapter 3

From social categories to stereotypes: Introducing the concept of stereotypes in research with adults

4.1. Introduction

Stereotypes serve similar functions as social categorical knowledge: They generate shared sets of expectancies about the likely traits, behaviors, preferences, or physical features of social category members (Dovidio, 1999; Over & McCall, 2018; Quadflieg & Macrae, 2011). As such, the processes involved in social categorization are assumed to be largely related to those that underly the development of stereotypes (Allport, 1954; D. Hamilton, 1981). The empirical research reviewed and presented in Chapter 1 and Chapter 2 provides evidence for children's early capacity to expect that category members share similar properties (Heyman & Gelman, 2000a, 2000b; Kalish, 2012; Moty & Rhodes, 2021; Riggs et al., 2014b), that the features marking category members will endure over time (Gelman, 2003; Liberman et al., 2017), and that category members ought to engage in prosocial behaviors toward one another and conform to the category's norm (Kaufmann & Clément, 2014; Over & McCall, 2018; Rhodes, 2013; Rhodes & Chalik, 2013). These capacities are important for the development of stereotypes in that they underlie the formation of strong and stable representations about social categories (S. J. Sherman et al., 2013). As children (and adults) start to identify with and attach meaning to particular social categories, social motives might intervene, driving the formation of biased representations about social categories, i.e., stereotypes, and differential evaluations of social categories, i.e., intergroup bias (Over & McCall, 2018; Shutts & Kalish, 2021).

The goal of this chapter is to delineate how we go from categorizing social agents to holding stereotypes. While the previous chapters of this dissertation offered a developmental perspective on social categorization, the following chapters present four experiments that assessed the effects of stereotypes on adults' information processing. An exhaustive review of the vast literature on stereotyping is beyond the scope of this chapter and would constitute on itself a whole book (as for e.g., Banaji & Greenwald, 2013; McGarty et al., 2002). Rather than a comprehensive review of theoretical perspectives and empirical findings, this chapter aims to

extract some basic principles that have emerged from social psychological research on stereotypes and to introduce key concepts that are relevant for the remaining chapters. Specifically, this chapter addresses the key question of how and why stereotyping differs from general categorical thinking by presenting cognitive bias and social motives that underly stereotypes (Section 4.2). It will then turn to the mental processes involved in stereotyping and present classical approaches to probe stereotype activation (Section 4.3), thereby introducing key methodological considerations for the empirical works presented in the second part of this dissertation (Section 4.4).

4.2. From social categories to stereotypes

The social categorization process establishes that some people share family resemblance and can thus be grouped together, and differentiated from others, based on their similarities (E. E. Smith & Medin, 1981). Once an individual has been ascribed to a social category, one can readily infer the likely properties of this individual. Stereotypes, however, go beyond this categorical knowledge in that they are biased, oversimplified, representations of social categories. One key question thus emerges: What are the conditions that lead to the emergence of stereotypes? Since the first time the term *stereotype* was introduced by Walter Lippmann a century ago (Lippmann, 1922), multiple theoretical perspectives have been proposed to account for the underlying factors responsible for the formation of stereotypes. One approach posited that cognitive bias and limited processing capacity were mainly responsible for the emergence of stereotypes (Simon, 1955; Stroebe & Insko, 1989; Tajfel, 1969). While these cognitive accounts (e.g., Simon, 1955; Tajfel, 1969) viewed the role of affective and motivational factors as minimal, other perspectives showed that the concept of stereotypes could not be accurately accounted for without considering fundamental social motives, such as needs to enhance self-esteem (e.g., Cvencek et al., 2021; Greenwald et al., 2002; Tajfel & Turner, 1979), to explain social differences (e.g., Jost & Hamilton, 2004; Yzerbyt et al., 1997, 2001), and to maintain power and social status (e.g., Fiske et al., 2002; Jost & Banaji, 1994; Jost & Kay, 2005). Addressing the key question of why stereotyping differs from general categorical thinking, the following sections review cognitive approaches to stereotypes before turning to theoretical perspectives focusing on fundamental social motives that underly stereotyping.

4.2.1. The cognitive approach

In the late 1960s, the cognitive approach (e.g., Simon, 1955; Tajfel, 1969) mainly conceived stereotypes as categories that structure and organize our social environment. According to the cognitive approach, the biased, oversimplified, representations of social categories, characteristic of stereotypes, resulted from limitations of the human capacity for processing information (Simon, 1955; Stroebe & Insko, 1989; Tajfel, 1969). When social agents are classified into categories, perceivers start to view members of the same social category as more homogenous, and members of different social categories as more different from each other (Eiser & Stroebe, 1972; Ford & Stangor, 1992a; Tajfel & Wilkes, 1963). On this account, stereotypes result from the exaggeration and overgeneralization of differences between groups and similarities within groups. These within- and between-group distortions tend to generalize to additional features, beyond those that differentiated the social categories in the first place (Allport, 1954; Gaertner & Dovidio, 2004), and these features are ascribed indiscriminately to *all* members of a social category (Jost & Hamilton, 2004).

Early empirical works on categorization provided evidence for this cognitive perspective, showing that when differences between groups already exist, these differences tend to be accentuated. Initiated by the works of Tajfel and Wilkes (1963), research on this topic revealed that the mere categorization of graded stimuli into discrete groups led to the exaggeration of perceived differences between categories and perceived similarities within categories. However, another important line of research started to investigate the possibility that stereotypes might emerge in the absence of group differences (Stroebe & Insko, 1989). Theories of illusory correlations (D. L. Hamilton & Gifford, 1976) notably proposed that stereotypes could form on a basic tendency to overestimate the frequency of co-occurrence of events which are statistically infrequent. To probe this assumption, Hamilton and Gifford's (1976) designed a paradigm in which participants were first presented with behaviors of social group members (e.g., "Bill, a member of group A, always talk about himself and his problems") and then, asked to guess the behavior of a new member from group A. The size of the groups and the frequency of the behaviors were manipulated, so that participants were exposed to groups of differing sizes and to behaviors that occurred more frequently than others. Across two experiments, Hamilton and Gifford (1976) found that participants associated the least common properties with minority groups (i.e., those composed of fewer members) and the most common properties with majority groups (i.e., those with more members). The effects of illusory correlations have, since then, been widely replicated (e.g., Acorn et al., 1988; D. L.

Hamilton et al., 1985; D. L. Hamilton & Rose, 1980; Sanbonmatsu et al., 1987), and it is now well established that observers might wrongly assume that a correlation exists between two distinct events that co-occur.

More recently, Sherman et al. (2009) developed a new model of stereotype formation based on the core principles of illusory correlations. Their model postulates that people first learn about the properties of frequent categories, simply because frequent categories are more numerous and more likely to be encountered than infrequent categories. When learning about rarer categories, perceivers focus their attention on the dimensions that best distinguish rare categories from the previously learned frequent categories. Because of this increased attention to dimensions that best distinguish social categories, people come to form a stronger association between infrequent categories and their properties, than between frequent categories and their properties (J. W. Sherman et al., 2009). An extension of this model further proposes that novel, rarer, social categories are more likely to be associated with negative attributes than with positive attributes (Alves et al., 2018). This proposition stems from research on the structure of the information ecology (Alves et al., 2017a, 2017b), showing that negative attributes are more diverse but less frequent than positive attributes. It follows that the attributes that best distinguish novel groups from known ones will be most likely negative. Importantly, both classical frameworks (Eiser & Stroebe, 1972; D. L. Hamilton & Gifford, 1976; Tajfel & Wilkes, 1963) and more recent models (Alves et al., 2018; J. W. Sherman et al., 2009) share the assumption that stereotypes are the result of information-processing biases. Collectively, cognitive approaches provided significant insights into the mental processes, i.e., accentuation effects and illusory correlations, responsible for stereotype formation. However, by confining stereotypes to cognitive biases, these accounts disregarded fundamental social motives involved in stereotyping.

4.2.2. Social motives underlying stereotypes

Because stereotypes refer to categories of people, they are self-relevant and socially relevant in ways that general categorical knowledge is not (S. J. Sherman et al., 2013). While categorizing others, we always also implicitly categorize ourselves, and come to divide the world into ingroups, the people to whom we belong, and outgroups, those to whom we do not belong (P. M. Brown & Turner, 2002; Dunham, 2018). The social group to which people belong becomes an important determinant of their social identity (Tajfel & Turner, 1985). This, in turn, motivates people to view their ingroup positively, which might be achieved by associating less

favorable properties with outgroup members compared to ingroup members (Dovidio, 1999; Judd & Park, 2004; Stroebe & Insko, 1989). Moreover, stereotypes might also be intertwined with essentialist beliefs when social categories are viewed as natural kinds. In this case, stereotypes go beyond categorical knowledge in that they provide explanations, intuitive theories, as to why category members belong together (Jost & Hamilton, 2004; Yzerbyt et al., 1997, 2001). Finally, stereotypes are also more than sets of descriptive properties about social categories, insofar as they might act as "justificatory device" (Allport, 1954, p. 192) that rationalize and perpetuate existing hierarchical structures in society (Jost et al., 2004; Jost & Banaji, 1994).

The following sections present the above three fundamental social motives that account for why stereotypes generate biased representations of social categories, namely because stereotypes serve to enhance self-esteem (Section 4.2.2.1), to explain social differences (Section 4.2.2.2), and to maintain power and social status (Section 4.2.2.3).

4.2.2.1. Stereotypes serve to enhance self-esteem

A rich line of research, starting with Tajfel (1970), has shown that people tend to favor those with whom they share even a slight, arbitrary feature (Dunham, 2018; Richter et al., 2016). These effects have been greatly examined under minimal group conditions. For instance, research involving Minimal Group Paradigms (Tajfel, 1970) showed that, when assigned to an abstract, meaningless group (e.g., those wearing blue shirts *versus* those wearing red shirts), people form more positive impressions of their ingroup compared to the outgroup, and evaluate more positively ingroup members than outgroup members (e.g., Howard & Rothbart, 1980; Otten & Wentura, 1999; J. W. Sherman et al., 1998; Tajfel et al., 1971; Van Bavel & Cunningham, 2009). These effects of mere membership have also been reported among children. From 5 years old, children prefer people with whom they share only a minimal similarity, distribute more resources to, and remember more positive information about ingroup compared to outgroup members (e.g., Baron & Dunham, 2015; Dunham et al., 2011; Plötner et al., 2015; Sparks et al., 2017; Wilks & Nielsen, 2018). According to the Social Identity Theory (Tajfel & Turner, 1979), effects of mere membership appear because of an underlying motivation to protect the ingroup's image, and at the same time, one's self-esteem. People with low self-esteem are prone to evaluate their ingroup more positively than the outgroup as a way to achieve a "boost" in their self-esteem (Tajfel et al., 1971). Alternatively, the Balance Identity Theory (Greenwald et al., 2002) postulates that sense of self-esteem is associated with the

ingroup such that a person with positive high self-esteem will evaluate more positively their ingroup than a person with low self-esteem. While these two theories differ in the way self-esteem is conceived, they both predict that when people identify with a group, the ingroup's welfare and image become closely related to one's own well-being and self-esteem (Cvencek et al., 2021; Greenwald et al., 2002; Tajfel & Turner, 1979).

The motivation to view one's own group positively plays a crucial role in the formation and content of stereotypes. Schaller and Maass (1989), for instance, ascribed participants to either a majority or a minority group and found that participants tended to see their own group more favorably than the outgroup, regardless of how the two groups were described. This is because the motivation to protect the ingroup's image tends to color in- and outgroup stereotypes. As almost any semantic content, especially trait dimensions, inevitably bears an evaluative component (Kurdi, Mann, et al., 2019; Peabody, 1967; Rosenberg & Sedlak, 1972), the properties assigned to a group can be interpretated at times as positive and at other times as negative (Campbell, 1967). For instance, the ingroup can be praised for being generous and ambitious, while an outgroup can similarly be devalued for being too generous or too ambitious. Stereotypes are, in this sense, biased representations of social categories, that can operate in the service of intergroup bias and prejudice (Judd & Park, 1993, 2004).

4.2.2.2. Stereotypes serve to explain social differences

Stereotypes are also distinct from categorical knowledge, insofar as they can provide explanations, intuitive theories, as to why category members belong together (Murphy & Medin, 1985). Although social categories are, in reality, artifactual, people might come to view social categories as natural kinds, and endow category members with some kind of essence, i.e., a common unalterable inner core (Diesendruck, 2020; Fiske & Taylor, 2013; Jost & Hamilton, 2004; Rothbart & Taylor, 1992). The category's essence mainly rests on erroneous interpretations of social groups, whereby group members are believed to belong together because of biological reasons, or something endogenous to the person, rather than because of social constructions that evolve with history and culture (Fiske & Taylor, 2013). Conceiving social categories as natural kinds has two important consequences. First, it licenses greater confidence in drawing inferences about category members. People perceive category members as being highly similar to one another, especially on properties related to the category's essence (Yzerbyt et al., 1997, 2001). This ultimately leads people to make sweeping generalizations about *all* category members (Allport, 1954; Haslam et al., 2002; Jost & Hamilton, 2004).

Second, category membership is viewed as unalterable and naturally determined (Gelman, 2003; Pauker et al., 2020). People thus tend to attribute differences between social categories to natural intrinsic causes, instead of to flexible societal factors (Rhodes et al., 2012; Rhodes, Leslie, Saunders, et al., 2018). As a consequence, when stereotypes are associated with essentialist beliefs, they go beyond the mere description of groups in terms of commonly associated properties to explain the origins of category membership (Jost & Hamilton, 2004). In this case, stereotypes become inflexible reductionist representations of social categories that ignore the complexity and diversity of people (Allport, 1954; Dovidio, 1999).

4.2.2.3. Stereotypes serve to maintain power and social status

Allport (1954) postulated that stereotypes are shared forms of justification for social differences (that often reveal to be erroneous). Forty years later, Jost and Banaji (1994) specified this claim, proposing that stereotypes serve three justificatory functions. The ego- and group-justification functions suggest that stereotypes fulfill people's need to perceive themselves and their ingroup favorably, and this need might also translate into devaluing others (Jost & Hamilton, 2004). The general idea behind ego- and group-justifications is mainly related to theories of stereotypes as a means to enhance self-esteem (cf. Section 4.2.2.1, this chapter). Jost and Banaji (1994) additionally noted that stereotypes might serve system-justification motives that justify the status quo and render cultural practices and institutions (e.g., segregation, capitalism, patriarchy) legitimate, rational, and even inevitable. This perspective holds that (a) members of disadvantaged as well as advantaged categories will engage in system justification to rationalize the status quo, (b) stereotypes justify and maintain particular social differences and divisions of labor, and (c) essentialist forms of stereotyping should emerge because they render intergroup differences natural and significant (Jost et al., 2004; Jost & Banaji, 1994; Jost & Hamilton, 2004).

Following the System Justification Theory (Jost & Banaji, 1994), members of both high- and low-status groups engage in thoughts and behaviors that reinforce and justify the existing social hierarchy. Empirical studies provide support for this proposition, showing that members of disadvantaged groups sometimes have evaluative preferences for outgroup members (e.g., Jost et al., 2002; Jost & Burgess, 2000; Newheiser et al., 2014). Moreover, members of disadvantaged groups have been reported to hold unfavorable stereotypes about their own group and positive stereotypes about high-status outgroups (e.g., Ashburn-Nardo & Johnson, 2008; Jost et al., 2004; Sniderman & Piazza, 1993). This outgroup favoritism is seen as one

manifestation of the tendency to internalize and perpetuate social inequalities (Jost et al., 2004). Furthermore, research focusing on structural relations between groups established that the contents of stereotypes might serve to sustain social hierarchies. The Stereotype Content Model (Fiske et al., 2002) for instance, posits that stereotypes are organized along warmth and competence dimensions according to the status and competence of one's ingroup in comparison to the other outgroups (Cuddy et al., 2009; Fiske et al., 2002). If a high-status outgroup is perceived as very competent, it is often also viewed as very low in warmth (e.g., competent but arrogant), whereas a low-status outgroup considered as incompetent is also seen as warm (e.g., poor but happy). These contrasting, yet complementary, stereotypes serve to maintain and justify the status quo, insofar as such stereotypes preserve the beliefs that all groups have some rewards, and no group possesses all valued characteristics (Jost et al., 2004; Jost & Hamilton, 2004; Kay et al., 2005; Kay & Jost, 2003). In this view, these complementary stereotypes confer legitimacy to the system as everyone benefits from it in at least some respects (Jost & Kay, 2005).

As presented in the previous section, stereotypes and essentialist beliefs can be intertwined, which drives people to view social group differentials as meaningful and inevitable. Moreover, if group members are *fundamentally* agentic, communal, lazy, ambitious, etc., then the group's position in social hierarchies is not only natural but also justified (Jost & Hamilton, 2004; Yzerbyt et al., 1997). In a recent empirical study, Mandalaywala et al. (2018) aimed to determine whether essentialism would increase endorsement of social hierarchies and existing inequalities. In this study, Black and White participants were exposed to fictional science news articles that either promoted essentialism (e.g., there exist genetic underpinnings of race) or not (e.g., race has no genetic basis). When essentialist beliefs were induced, both White and Black participants showed greater endorsement of social hierarchies (as measured by Social Dominance Orientation-6 Scale; Pratto et al., 1994). That is, the extent to which people essentialized social categories promoted the beliefs that social differentials reflected objective and natural structures in the world (Mandalaywala et al., 2018; Yzerbyt et al., 1997). In this case, stereotypes act as "justificatory device" (Allport, 1954, p. 192) that rationalize and perpetuate existing hierarchical structures in society[11].

[11] Please note that both high- and low-status groups are likely to portray outgroup members as essentially different from ingroup members. Dominant groups might engage in essentialist thinking to justify and maintain their privileged status (Diesendruck, 2020, 2021; Mahalingam, 2003; Sidanius & Pratto, 1999), while minority groups might recruit essentialism to form a unified force and bolster group identity (Mandalaywala et al., 2018; Morton & Postmes, 2009; Ryazanov & Christenfeld, 2018).

Summary

The conclusion to be drawn from the reviewed accounts of stereotyping is that stereotypes are more than sets of descriptive properties associated with category members. Stereotypes reflect biased oversimplified representations of social categories that can operate in the service of intergroup bias and justify social inequalities. While the cognitive approach proposed that stereotypes mainly emerge from limitations of the human capacity for processing information, other theoretical perspectives insisted on the predominant role of social motives in the emergence of stereotypes. To this point, the present chapter shed light on the emergence and functions of stereotypes by delineating cognitive bias and social motives involved in stereotyping. The next section outlines the mental processes involved in stereotyping (Section 4.3), before introducing traditional measures of implicit social cognition (Section 4.3.1) and examining some key issues currently under debate concerning these measures (Section 4.3.2). The chapter concludes with a presentation of reading tasks as a promising additional way to probe implicit processes involved in stereotyping (Section 4.4), thereby providing key methodological considerations for the empirical works presented in the second part of this dissertation.

4.3. Underlying mental processes involved in stereotyping

Starting from the 1980s, research in social psychology provided convincing evidence that stereotyping processes could operate in unconscious and unintentional ways, as well as consciously. This dual view of stereotypes, as automatic and controlled, builds on classical dual-process conceptions of mental operations. Dual-process theories share the generic idea that the mind operates in rather automatic or controlled manners, where automatic processes, unlike controlled processes, occur without effort, awareness or intention and require little cognitive resources[12] (Blair, 2002; Fiske & Taylor, 2013; Greenwald & Banaji, 2017). In particular, associative networks models (Anderson & Bower, 1973; Collins & Loftus, 1975; Wyer & Carlston, 2018) posit that the activation of an item stored in memory travels through associated links to activate related items. The stronger the association, the more accessible the items are. In this perspective, learning is a critical component for automaticity. Following early

[12] To keep this section focused, a full description of all dual-process models is avoided. This choice is motivated by the fact that nearly thirty models of dual-process have been introduced over the last fifty years in various domain of research, such as person perception, causal attributions, or attitudes (Fiske & Taylor, 2013; Stanovich et al., 2014).

works in cognitive psychology (e.g., Posner & Snyder, 1975; Shiffrin & Schneider, 1977), well-learned items are automatically and easily retrieved from memory, while new weakly learned items require increased cognitive effort to be recalled. This model constituted an important conceptual root for the understanding of the processes involved in stereotyping.

Stereotypes are forms of information that have been – unintentionally – acquired over many years from language and experiences that "cumulatively construct an overlearned repertoire of cultural expertise" about social categories[13] (Greenwald & Banaji, 2017, p. 866). Stereotypes are thus conceived as well-learned items stored in semantic memory in a dormant state until they are activated in a given situation (Devine, 1989; Gilbert & Hixon, 1991; Macrae & Bodenhausen, 2000). The activation of stereotypes is considered as automatic because it happens relatively quickly and outside of awareness: In the presence of a category member (e.g., upon encountering a category member, when referring to a social category in a conversation, etc.), a set of related, possibly erroneous, concepts activates automatically in people's minds (Gilbert & Hixon, 1991; Krieglmeyer & Sherman, 2012; Quadflieg & Macrae, 2011). Under some conditions, such as ego-protective motivations (Fein & Spencer, 1997), personal tendency to hold essentialist beliefs (Levy et al., 1998), or contextual and cognitive demands (van Knippenberg et al., 1999), people might use the activated stereotypical expectations to judge or evaluate others (Macrae & Bodenhausen, 2000; McGarty, 2002; Quadflieg & Macrae, 2011). The term *implicit stereotypes* traditionally refers to the first process, i.e., the automatic activation of associations about social categories, whereas the term *explicit stereotypes* reflects the application of stereotypes, i.e., consciously endorsed beliefs about social categories.

Early research on stereotypes, mainly relied on self-report measures, such as asking participants to rate category members on semantic differentials scales[14] (e.g., Greenwald et al., 1998), or on explicit judgements tasks, such as asking participants whether they would hire a candidate for a stereotype-related job (e.g., Glick et al., 1988). These tasks thus traditionally assessed explicit stereotypes. However, research on attitudes and stereotypes started to show that these biases were more often subtle, automatic, and unintentional, than controlled and

[13] Please note that recent empirical studies also revealed that people might form implicit attitudes or stereotypes from brief exposure to evaluative statements (e.g., "Longfaces are bad"; Charlesworth et al., 2020; Gonzalez et al., 2017; Lane et al., 2020, see also Chapter 1, section 2.3.4, this dissertation).

[14] Semantic differential scales traditionally present participants with a concept (e.g., woman) and a list of opposite adjective pairs (e.g., strong-weak, caring-distant). Participants have to indicate on the scale whether the concept is more or less relevant with one of the paired adjectives (Greenwald et al., 1998).

blatantly expressed (Fiske & Taylor, 2013; Kurdi & Banaji, 2022). Diverse indicators revealed that implicit stereotypes and prejudices persisted, even though openly negative attitudes and explicit stereotypes started to dissipate (Crosby et al., 1980; Saucier et al., 2005). For this reason, the utility of self-report measures was questioned, insofar as they might fail to accurately capture implicit stereotypes because people tend to misreport their stereotypes because of social desirability (Blair, 2002) or because they underestimate or might even be unaware of their automatic associations (Nosek et al., 2011). This pushed social psychologists to consider other methods to assess implicit stereotypes. Inspired by research on implicit memory and semantic association (e.g., Collins & Loftus, 1975; Jacoby, 1991; Neely, 1976; Schacter & Graf, 1986), social psychologists started to examine automatic processes in the domain stereotypes and attitudes with implicit measures, opening the way to a flourishing era of research on implicit social cognition.

4.3.1. Measuring implicit forms of stereotypes

Social psychologists in the mid 1980s started to employ methods from research on associative memory and semantic priming in the domain of attitudes and stereotypes (Anderson & Bower, 1973; Collins & Loftus, 1975; Wyer & Carlston, 2018). Adapting semantic priming methods (e.g., Meyer & Schvaneveldt, 1971; Neely, 1976), Gaertner and McLaughlin (1983) offered one of the first assessment of the automatic activation of racial stereotypes with implicit measures. In their study, participants were instructed to decide, as fast and as accurately as possible, if letter strings were words or not. The critical manipulation was that the pairs either matched racial stereotypes (e.g., smart-White, lazy-Black) or not (e.g., smart-Black, lazy-White). Their findings showed that participants were significantly faster to identify words when the pairs were consistent with racial stereotypes than when they were inconsistent with racial stereotypes. A remarkable proliferation of research assessing the automatic activation of stereotypes with implicit measures soon followed. Fazio et al. (1986) notably developed the now famous procedure of evaluative priming – also an adaptation of semantic priming – to assess the automatic activation of attitudes. In Fazio et al.'s (1986) procedure, participants categorized target words (e.g., delightful, repulsive) as pleasant or unpleasant after being presented with a valanced prime (e.g., music, crime). Participants were faster to classify a target word as pleasant when it followed a positive prime than when it followed a negative prime. This procedure has then been extensively used, notably to assess automatic attitudes and

stereotyping about various social categories (e.g., Banaji & Hardin, 1996; Bargh & Pietromonaco, 1982; Devine, 1989; Dovidio et al., 1986; Wittenbrink et al., 1997).

Based on those earlier procedures, Greenwald and Banaji (1995) developed the Implicit Association Test (IAT; Greenwald et al., 1998). As its name suggests, the IAT provides a measure of relative associations strengths between concepts (Greenwald et al., 1998). However, the IAT differed from the previous evaluative priming tasks in that it did not involve priming procedures. Rather, the IAT consists in asking participants to classify stimuli in four different categories with just two response options (Greenwald & Banaji, 2017). Concretely, in IAT procedures, a pair of concepts (e.g., women/men, elderly/young people, flowers/insects) is combined with a pair of attributes (e.g., home/career, weak/strong, pleasant/unpleasant) before reversing the association. The critical measure in IAT is the time participants need to sort the stimuli in the four categories (e.g., women, men, home, career). Pairs of categories eliciting faster responding are assumed to be more strongly associated than those eliciting slower responding (Greenwald et al., 2003). In other words, the IAT provides a measure of mental effort required to give the same response to two non-associated categories (e.g., men and home) (Greenwald & Banaji, 2017). Both evaluative priming tasks and IATs were qualified as implicit to the extent that they assess psychological attributes (e.g., attitudes, stereotypes) without involving explicit self-report (Fazio & Olson, 2003; Gawronski et al., 2020). Today, implicit measures have been widely employed to assess automatic processes underlying judgments and social behavior (for recent reviews, see Gawronski et al., 2020; Greenwald & Lai, 2020). Yet, implicit measures are also at the center of an active debate and raised important theoretical controversies and methodological questions.

4.3.2. Ongoing debates on implicit measures

One broad issue in the literature on implicit measures pertains to the proper use of terminology, in particular regarding the term *implicit*, which has been used to describe both measurement procedures and assessed psychological constructs (Payne & Gawronski, 2010, see also Greenwald & Banaji, 2017; Greenwald & Lai, 2020; and the *Handbook of Implicit Social Cognition* edited by Gawronski & Payne, 2010). These competing usages of the term *implicit* date back to early research on memory. Influential works on memory with indirect measures (e.g., Graf & Schacter, 1985; Jacoby & Dallas, 1981; Jacoby & Witherspoon, 1982) produced evidence for memories that were not apparent on direct measures (e.g., free recall or recognition tasks). Following these findings, the concept of *implicit memory* was introduced

and defined as traces of past experience that influence later performance, in the absence of *conscious* recollection for the earlier experience (Schacter, 1987). However, this perspective was soon questioned by the works of Jacoby (1991) and Reingold & Merikle (1988) who concluded that both direct and indirect measures combined influences of conscious *and* unconscious mental processes. Similarly, in the domain of implicit social cognition, some researchers used the term *implicit* to refer only to *indirect* measurement procedures, which assess automatic associations without involving self-reports (de Houwer & Moors, 2010; Greenwald & Banaji, 2017). Different from this view, other researchers sometimes used the term *implicit* to describe "unconscious representations that are inaccessible to introspection" (Gawronski et al., 2020, p. 3). This latter perspective generated many controversies over the last 25 years (e.g., Gawronski et al., 2020; Greenwald & Banaji, 2017; Greenwald & Lai, 2020) because it is not possible to consider implicit measures as pure indicators of unconscious mental processes. To date, no clear conceptualization of what the term *implicit* entails, i.e., automatic or unconscious processes, has reached the consensus and researchers in the domain advocate for the (consensual) empirical meaning of the term *implicit* as referring to indirect measures of social constructs, "without implication for the conscious or unconscious nature of what those measures reveal." (Greenwald & Lai, 2020, p. 421; De Houwer et al., 2009; Fazio & Olson, 2003; Greenwald & Banaji, 2017) However, this position leaves an important conceptual issue unresolved: What are the features, i.e., conscious/unconscious, controlled/automatic, intentional/unintentional, of the psychological constructs that are assessed by indirect measures?

Another key issue pertains to the relationship between implicit measures and behavioral outcomes. Over the years, implicit measures gained increased interest both in academic and public domains, especially because they offer a potential explanation for why group-based inequalities remain, while explicit intergroup attitudes and stereotypes tend to dissipate (Charlesworth & Banaji, 2019). An important number of research thus started to examine the relationship between implicit measures – with a focus on IAT performance – and intergroup behaviors. In 2009, a first meta-analysis was conducted by Greenwald and colleagues who observed that implicit measures (i.e., IAT scores) had greater predictive validity than explicit self-reports regarding socially sensitive topics (e.g., intergroup discrimination) (Greenwald et al., 2009). However, their meta-analysis also revealed that in seven of the nine criterion domains examined (e.g., sexual orientation preferences, psychological health, political preferences) explicit measures outperformed the IAT in the prediction of behavior. These results led to a

subsequent meta-analysis conducted by Oswald et al., (2013) who concluded that implicit measures were no better predictor of behavior than explicit measures, reporting smaller predictive validity correlations than Greenwald et al. (2009). Almost ten years later, a large meta-analysis (Kurdi, Seitchik, et al., 2019), using data from more than 200 of research, showed that the average predictive validity correlations of both explicit *and* implicit measures were small but nonetheless statistically significant and almost identical in size (Kurdi, Seitchik, et al., 2019; Kurdi & Banaji, 2022). These results provided convincing evidence that would settle the debate, but one important theoretical issue remains open: How exactly does implicit cognition drive behavior[15]?

4.4. Conclusion and remarks on the following empirical studies

In the last 25 years, a growing number of research started to use indirect measures to assess implicit attitudes and stereotypes. Unlike explicit self-report measures, indirect measures of attitudes and stereotypes rely on less controllable behaviors, such as response times, error rates or other responses that are likely to bypass deliberate control. Although indirect measures are still a subject of ongoing research and debate within the field of social psychology, they nonetheless provide valuable insights into subtle automatic forms of stereotyping. Indeed, indirect measures solve the problem that people might misreport their stereotypes because of social desirability (Blair, 2002) or because they underestimate or might even be unaware of their automatic associations (Nosek et al., 2011). Probing the automatic activation of stereotypes with indirect measures thus offer important insights into why stereotypes and prejudices persist, despite growing effort and attention to eliminate them. The following empirical works build on this research tradition, assessing with reading tasks how stereotype activation affects information processing.

As presented earlier in this chapter, stereotypes operate in automatic and unintentional ways, as well as explicitly and deliberately. One theoretical perspective posits that the underlying processes involved in stereotypes rest on associative networks models (Anderson & Bower, 1973; Collins & Loftus, 1975; Wyer & Carlston, 2018). Accordingly, the activation of

[15] One key result emerging from Kurdi, Seitchik, et al.'s (2019) meta-analysis suggests that implicit cognition might first influence explicit cognition which, in turn, drives behavior (Greenwald & Banaji, 2017). Yet, this proposition is contrary to dominant dual-process theories advocating for a dissociation between implicit and explicit measures, i.e., implicit measures capture automatic behaviors, whereas explicit measures capture controlled behaviors (Gawronski et al., 2006, 2020).

a concept (e.g., woman) travels through associated links to activate related concepts (e.g., caring, delicate, mother). The stronger the association, the easier it is to process related concepts. In this sense, stereotypes are thought of as networks of related attributes about social categories that are stored in semantic memory in a dormant state until they are activated in a given situation (D. E. Carlston, 1992; Gilbert & Hixon, 1991; Krieglmeyer & Sherman, 2012; Quadflieg & Macrae, 2011). Classical indirect measures of stereotypes typically assess whether words or images of social categories (e.g., men, women) facilitate the processing of semantically congruent concepts (e.g., career, home). The relative automatic activation of social categories and related attributes is inferred by comparing differences in response latency or accuracy across stereotype-congruent and stereotype-incongruent trials (Fazio & Olson, 2003; Greenwald et al., 1998; Kurdi, Seitchik, et al., 2019). The methodology employed in the next empirical studies follows the same principles.

Two reading tasks (a self-paced reading and an eye-tracking tasks) were created to capture the automatic activation of stereotypes. The experimental design consisted in presenting sentences whose content either matched or violated stereotypical expectations. The central empirical question was to investigate if reading about category members (e.g., Italians) automatically activated stereotypical expectancies about the likely attributes of this social category (e.g., being seducers). Similar to classical indirect measures' assumptions, faster reading times on sentences that contained stereotype-congruent information (such as being seducers) than on sentences that contained stereotype-incongruent information (such as being shy) were taken as evidence for the automatic activation of stereotypes. The methodology employed in the next studies, however, differs somewhat from the standard ways of assessing implicit stereotypes. Unlike classical indirect measures, reading tasks do not resort to sorting or priming procedures, offering an interesting asset to the study of stereotype activation because they are close to natural settings and do not interrupt the ordinary flow of cognition. Moreover, the sentences used in the following studies were based on extracts found in in movies, hearsay, comics, magazines, etc., matching stereotypical contents to which people might be exposed in their daily life. Reading tasks using real-world extracts therefore have the potential to reflect natural reactions to such stimuli.

On a final note, the approach that is taken in the following studies is to adopt the definition of *implicit* as referring to *indirect* measurement procedures, which assess automatic associations without involving self-report (de Houwer & Moors, 2010; Greenwald & Banaji, 2017). It makes no assumption regarding the implicit nature of the psychological constructs

under investigation. Participants in the following studies might have been aware of the stereotypes involved when performing the tasks, but were unlikely to control the effects of stereotype activation because of the employed measurement procedures.

5. Chapter 4

Social cognition and Relevance: How stereotypes impact the processing of definite and indefinite descriptions[16]

Abstract

This paper focuses on the impact of social cognition on the processing of linguistic information. More specifically, it brings some insights to Relevance theory's construal of Meaning$_{NN}$ (Sperber & Wilson, 2015), which seeks to account for non-propositional meanings (D. Wilson, 2016). It shows, through two experiments, how gender and nationality-related stereotypes guide the processing of definite and indefinite descriptions. Experiment 1 consists of a self-paced reading task (with 59 French native speakers), introducing information confirming *versus* violating gender stereotypes within a nominal phrase (NP). The NP (e.g., "chirurgien/chirurgienne", "surgeon$_{male/female}$") was itself introduced either by a definite article (presupposition) or an indefinite article (assertion). Results showed that information violating gender stereotypes was costlier to process than stereotype-congruent information. Moreover, when information violated gender stereotypes, definite descriptions became significantly costlier than indefinite ones, because they required the identification of a salient referent which contradicted stereotypical expectations. Experiment 2 tested the effects of definite *versus* indefinite NP on processing nationality-related stereotypes in a self-paced reading task (with 49 French native speakers). Participants read definite *versus* indefinite NPs referring to representatives of a country. The NP was subsequently paired with information that confirmed *versus* contradicted nationality stereotypes. Results showed that information contradicting nationality stereotypes were significantly costlier to process than information confirming stereotypes. Furthermore, when information contradicted nationality stereotypes, indefinite descriptions (which promote a single occurrence reading) failed to facilitate information processing compared to definite descriptions (which promote a generalized representation of the social category). Overall, the present findings are consistent with research on stereotypes,

[16] Reprint of: Mari, M. A., & Müller, M.-L. (2023). Social cognition and Relevance: How stereotypes impact the processing of definite and indefinite descriptions. *Frontiers in Communication, 8,* p. 1-16. https://doi.org/10.3389/fcomm.2023.1088861

This research was funded by the University of Neuchâtel.

in that they show that stereotype-incongruent information affect sentence processing. Importantly, while Experiment 1 revealed that stereotypes affected the processing of linguistic markers, Experiment 2 suggested that linguistic markers could not modulate the processing of stereotypes.

5.1. Introduction

This paper lies at the intersection of social cognition and pragmatics. Using tools from the study of stereotypes, it contributes to the theoretical framework of Relevance theory (Sperber & Wilson, 1986, 2015). Relevance theory argues that meaning derivation is guided by a comprehension heuristic. When exposed to an ostensive verbal stimulus, the listener seeks for optimal relevance, minimizing processing costs, to obtain most cognitive effects through the acquisition, reinforcement, or revision of a belief (cf. Sperber & Wilson, 2015, p. 135).

Initially, Relevance theory developed Grice's theory of implicatures (Grice, 1957, 1975), providing a cognitive explanation for pragmatic inferences responsible for explicit and implicit meanings. However, in recent years, it focused more on argumentation and literary studies (Cave & Wilson, 2018; Mercier & Sperber, 2011; Sperber et al., 2010; Wharton & Strey, 2019). In this context, Relevance theory presented a new research agenda oriented towards a broader approach to ostensive communication (Sperber & Wilson, 2015): it is emphasized that an adequate theory of meaning should include not only "determinate propositions" conveyed by linguistic stimuli, but also non-propositional meanings conveyed by verbal and nonverbal cues. Among the examples mentioned, they present the following exchange, for which the levels of analysis are broader than those initially proposed in classical approaches in pragmatics:

(1) Rob: Do you live in London?
Jen: I live in Chelsea

(Sperber & Wilson, 2015, p. 144)

In the above, Jen implicitly answers Rob's question in the affirmative, given that Chelsea is a neighborhood in London. However, the relevance of the utterance will not only depend on determinate contents (on the level of explicature or implicature), but also on less determinate ones, triggered by the tone of voice or the social status of the speakers (Sperber & Wilson, 2015, p. 144). Here, Sperber and Wilson point out that when social status is manifest, they will guide

inferences in different ways. For example, depending on their respective social status, Jen's utterance may express closeness because she shares more specific information about where she lives, or it may express a sense of social superiority that can be paraphrased as "I don't live in just any part of London."

According to Relevance theory, social status corresponds to "encyclopedic information", such as gender or nationality stereotypes in this study. Encyclopedic information is used by addressees to construct the context which guides them in making interpretive inferences. In response to an ostensive stimulus, the recipient constructs contextual hypotheses on the basis of information that is more or less salient, and respectively more or less easy to process. The construction of the context will allow the addressee to infer the premises leading to the derivation of an intentional explicature or implicature intended by the speaker (Sperber & Wilson, 1986, p. 37).

Furthermore, it should be noted that the most recent lines of research in Relevance theory argue that the comprehension heuristic should be conceived as a broader process than initially defined, accounting for less determinate meanings and including nonverbal cues (Sperber & Wilson, 2015, p. 137). Following these new perspectives, Wilson (2016, p. 15) argues that linguistic markers may activate clusters of domain-specific modules of cognition, such as mindreading, emotion reading, or social cognition. The activation of these domain-specific modules is presumed to have an effect on the relevance-guided comprehension heuristic.

The present study aims to contribute to current discussions in Relevance theory by testing the impact of gender and nationality-related stereotypes on the processing of specific linguistic information, namely definite and indefinite descriptions.

5.1.1. The processing of stereotypes

While reading, one must not only visually process the written words but also understand their underlying meaning. To comprehend a text, readers draw on different sources of knowledge, namely linguistic, orthographic, and general world knowledge (Kendeou et al., 2016; Perfetti & Stafura, 2014). As pointed out by Relevance theory (cf. example (1)), making inferences lies at the core of comprehension. With respect to reading tasks: readers retrieve information from memory to construct a mental representation of a text (Elbro & Buch-Iversen,

2013; Graesser et al., 1994; Kendeou et al., 2016). The mental representation combines elements that are derived explicitly from the text, as well as elements that are implicit, coming from the readers' previously acquired knowledge (Gygax et al., 2021). As such, readers' world knowledge plays an essential role in reading comprehension.

While reading a word or a sentence, related concepts are automatically activated in semantic memory (Gerrig & McKoon, 1998; O'Brien et al., 1998; Rapp & van den Broek, 2005; Rubio-Fernández, 2013). For instance, upon reading sentence (2), concepts such as LAWYER*, LAW COURT*, or CRIMINAL* are likely to be activated and be more accessible in readers' memory:

(2) The judge sentenced a burglar to two years in prison.

Similarly, theoretical accounts of stereotyping propose that a given situation might increase the accessibility of stereotypic knowledge in memory (Gilbert & Hixon, 1991; Quadflieg & Macrae, 2011; Rees et al., 2020). For instance, upon reading sentence (2), stereotypical representations of "the judge" and "a burglar" will be activated in readers' memory, allowing readers to hold expectations about the likely traits, features, and behaviors of the two protagonists (Beukeboom & Burgers, 2019; O. Klein & Bernard, 2015). From this perspective, stereotypes function as heuristics as they guide expectations about members of a social category and are rapidly processed (Krieglmeyer & Sherman, 2012; F. Müller & Rothermund, 2014). When information violates a stereotype, more cognitive effort is required to access stereotype-incongruent information from associative memory, leading to increased processing difficulty (Banaji & Hardin, 1996a; Bartholow & Dickter, 2008; Kutas & Federmeier, 2000; White et al., 2009). Importantly, previous research showed that regardless of personal opinions, people in the same context tend to be knowledgeable about the stereotypes in their culture (Beukeboom & Burgers, 2019; Devine, 1989; Lepore & Brown, 1997; Moskowitz et al., 1999; Quadflieg & Macrae, 2011). As such, if a word or a sentence refers to a social category, readers within the same culture will spontaneously produce inferences about this social category and will most likely hold similar stereotypical expectations.

Overall, the effects of stereotype information on reading are well documented. An important line of research assessed how gender stereotypes affect anaphora resolution of personal or reflexive pronouns (see for e.g., Carreiras et al., 1996; Duffy & Keir, 2004; Esaulova et al., 2014; Irmen, 2007; Kennison & Trofe, 2003; Reali et al., 2015). These studies

showed that reading times of anaphoric pronouns were longer when stereotypical expectations about role nouns did not match the gender of the pronoun (e.g., "The *firefighter* burned *herself* while rescuing victims from the building", Duffy & Keir, 2004, p. 553). Another line of research tested whether readers make inferences about the gender of a person upon reading a role noun (and so, not only when required by the anaphora). For instance, Garnham et al. (2002) designed a study in which readers could make inferences about the gender of a character, without involving anaphora resolution (e.g., "The soldier drove to the playgroup after work, and picked up one of the children, who said "Look what I did today daddy!"", Garnham et al., 2002, p. 442; see also Lassonde, 2015; Reynolds et al., 2006). Their findings suggest that readers automatically encode gender when they are exposed to role nouns, even though gender information is not crucial for comprehension (Gygax et al., 2021). Altogether, past studies revealed that, in various languages and cultures, reading is slowed down when *gender* is incongruent with stereotypical representations of role nouns (e.g., in English: Garnham et al., 2002; Lassonde, 2015; Reynolds et al., 2006; in Norwegian: Gabriel et al., 2017; in German: Esaulova et al., 2014; Irmen, 2007; in Italian: Cacciari et al., 1997; in Spanish: Carreiras et al., 1996). However, because gender is considered as a primary social category[17] (Brewer, 1988; Fiske, 1998), it is not clear whether information processing would be affected by stereotypical expectations about other social categories, that are less primary than gender, such as nationality stereotypes. For this reason, the present study compared, in a first experiment, the effects of well-studied primary stereotypes (i.e., professions associated with gender). In a second experiment, we assessed whether the observed effects also apply to less studied stereotypes, such as nationality stereotypes. Both experiments were designed to assess the extent to which stereotypes impact the processing of specific linguistic information (see next section below).

5.1.2. Research question and hypotheses

This study builds on Singh et al.'s (2016) experiment testing the impact of plausible *versus* implausible contexts on the processing of definite and indefinite descriptions. While definite descriptions trigger a presupposition of a salient referent, indefinite descriptions merely introduce a new referent (Singh et al., 2016, p. 619; but see also Sperber & Wilson, 1986, p. 706).

[17] Gender is considered as a primary social category because attention to gender emerges early (see for e.g., Quinn et al. 2002) and because children of 3-4 years of age are already aware of conventional gender stereotypes (see for e.g., Leinbach et al., 1997; Shutts et al., 2009; Weinraub et al., 1984).

According to Relevance theory, presuppositions and assertions can be distinguished in terms of foreground and background implications. While asserted contents contribute to relevance by providing additional cognitive effects, presuppositions contribute to relevance by saving efforts (Sperber & Wilson, 1986, p. 706). With respect to indefinite descriptions, they will be responsible for generating more effort because they present a noun as a new referent to the reader. This is not the case with definite descriptions, which present the noun phrase as "familiar" in context (cf. Heim, 1982; Roberts, 2003; Schwarz, 2009)

In Experiment 1, definite descriptions occur in a context that requires a bridging inference (Clark, 1975). That is, the referent is not explicitly mentioned in the preceding context, thereby requiring the construction of a link between the context (e.g., a hospital) and the noun (e.g., a/the surgeon). However, each context sentence was designed to have a strong semantic proximity with the target definite description, which facilitates processing (C. Clifton, 2013; Garrod & Sanford, 1977; Haviland & Clark, 1974; Schwarz, 2019). Bridging inferences are even easier in Experiment 2, as they involve a context introducing a superordinate concept (i.e., the name of a country), followed by a noun for a subordinate concept (i.e., the inhabitants of the country).

With regard to Singh et al.'s (2016) study, they hypothesized that implausible contexts, as in (3) below, would lead to an increased processing difficulty upon reading the following sentence. Moreover, within the implausible condition in (3), the definite description was expected to be significantly more difficult to process than the indefinite one, as it requires the identification of a salient referent in an incompatible context.

Singh and colleagues used two methods to test participants, namely a self-paced-reading task and a stop-making-sense task[18]. In both methods, participants read a plausible *versus* implausible context sentence (3), followed by a definite or indefinite noun phrase (henceforth NP). Implausible contexts were expected to make the target NPs significantly costlier than plausible ones. Furthermore, as mentioned above, definite NPs like (3b) were expected to be costlier than indefinite ones like (3a) within implausible contexts:

(3) Mary went to the beach$_{plausible}$ / office$_{implausible}$ a few hours ago.

(3a) A lifeguard warned her there about the weather.

[18] In the stop-making-sense task, participants were instructed to continue making words appear, segment by segment, as long as the sentences made sense. As soon as an incoming word or phrase did not make sense in the context of the preceding words/phrases, participants were asked to end the task (cf. Singh et al., 2016, p. 615).

(3b) The lifeguard warned her there about the weather.

(Singh et al., 2016, p. 631)

Singh and colleagues observed an effect of context plausibility, where implausible contexts made the target NP (*A/The* lifeguard) significantly costlier to process than plausible ones. However, no significant difference was found when comparing definite NPs with indefinite ones. An effect was only found in the stop-making-sense task, when summing over all participants: In this case the proportion of dropouts was significantly higher in the presupposition condition (*The* lifeguard) than in the assertion condition (*A* lifeguard) (Singh et al., 2016, p. 617). Importantly, no effect was found between presupposition and assertion conditions in the self-paced reading task (Singh et al., 2016, p. 618). In a replication of Singh and colleagues' study, using eye-tracking and self-paced-reading tasks, Müller and Mari (2021) found significant results for plausibility effects, but no difference between definite and indefinite articles in the implausible condition, just like Singh and colleagues.

The present study seeks to take these experiments further, using congruent *versus* incongruent stereotypes instead of plausible *versus* implausible contexts. The use of stereotypes, instead of context plausibility, is beneficial on two levels. First, it solves the problem of "context plausibility," which involves effects from various possible sources (e.g., surprise, comprehension problems, or also typicality effects). Importantly, the stimuli in this experiment used only plausible contexts, thus allowing the critical variable to be isolated, excluding surprise effects or problems attributable to the comprehension of the utterance. Second, as presented in the previous section, stereotypes are widely studied and well understood in terms of reading tasks.

Experiment 1 consisted in a self-paced-reading task, assessing the impact of gender stereotypes (i.e., a primary social category) on the processing of asserted *versus* presupposed contents. More specifically, Experiment 1 aimed to replicate previous findings on the effects of gender stereotypes on reading times cross-linguistically (with French speaking Swiss participants) and sought to identify the specific time course of processing gender stereotypic information. To this end, Experiment 1 tested the following hypothesis:

Hypothesis 1. Information violating gender stereotypes (4a) would be costlier to process than stereotype-congruent information (4b), within a compatible context (4).

(4) Lucienne est allée à l'hôpital le mois dernier.
(Context sentence)

(4a) *La/Une chirurgienne* l'a opérée avec une grande précision.
(Stereotype-incongruent)

(4b) *Le/Un chirurgien* l'a opérée avec une grande précision.
(Stereotype-congruent)

[Lucienne went to the hospital last month. (Context sentence)

The/A surgeon$_{female}$ operated on her with great precision. (Stereotype-incongruent)

The/A surgeon$_{male}$ operated on her with great precision. (Stereotype-congruent)]

Furthermore, and as in Singh et al. (2016) and Müller and Mari (2021), Experiment 1 tested whether definite NPs would lead to longer processing compared to indefinite NPs when the information contradicts a gender stereotype. In this case, the identification of a salient referent, required for definite NPs, is inconsistent with the encoding of stereotype-incongruent information.

Hypothesis 2. Stereotype-incongruent NPs would be costlier to process when presupposed through a definite description (e.g., "la chirurgienne*";* "the surgeon$_{female}$") than when asserted through an indefinite description (e.g., "une chirurgienne"; "a surgeon$_{female}$").

Experiment 2 focused on the processing of nationality-related stereotypes, i.e., a secondary social category, and their interaction with definite and indefinite descriptions. To our knowledge, only two papers have studied the processing of secondary social categories. Dickinson (2011) focused on stereotypical inferences regarding heterosexuality during reading tasks, and Lassonde (2015) assessed stereotypical expectations regarding the behaviors of social groups[19]. Whereas Lassonde (2015) found that reading times were longer for information that violated stereotypical expectations about social groups, Dickinson (2011) failed to reach conclusive results. Thus, given the limited information available on secondary social categories, it is worth providing new investigations.

In Experiment 2, participants first read a context sentence introducing the name of a country. Two countries were alternatively presented, for example Italy *versus* Japan, as

[19] In Lassonde's (2015) study, the stereotype-incongruent information was introduced by a whole sentence (e.g., "The nuns said there was not enough alcohol" *versus* "The rockers said there was not enough alcohol", (Lassonde, 2015, p. 161)). In Dickinson's (2011) study, the stereotype-incongruent information was initiated by anaphora resolution (e.g., "Last night, in the packed movie theatre Hannah screamed loudly until her wife held her close", (Dickinson, 2011, p. 457)).

presented below (5). The second sentence introduced a redundant NP ("A/The Italian/s" *versus* "A/The Japanese"), followed by an attribute ("great seducer/s") which was congruent (5a) or incongruent (5b) with a stereotype:

(5) Mathilde est allée en Italie / au Japon le week-end dernier.
(Context sentence)

(5a) Un/Les italien/s *a/ont joué au/x grand/s séducteur/s* durant tout le séjour.
(Stereotype-congruent)

(5b) Un/Les japonais *a/ont joué au/x grand/s séducteur/s* durant tout le séjour.
(Stereotype-incongruent)

[Mathilde went to Italy/Japan last weekend. (Context sentence)
An/The Italian/s *played the great seducer/s* during the whole stay. (Stereotype-congruent)
A/The Japanese *played the great seducer/s* during the whole stay. (Stereotype-incongruent)]

Theoretical perspectives on stereotyping propose that any kind of stereotype-incongruent information should be difficult to process because it requires more cognitive effort to access this information from associative memory (see for e.g., (see for e.g., Banaji & Hardin, 1996; Bartholow & Dickter, 2008; Kutas & Federmeier, 2000; White et al., 2009). Drawing from this perspective, the following hypothesis was tested:

Hypothesis 3. Information violating nationality stereotypes should elicit longer reading times than stereotype-congruent information.

As illustrated above, the noun introducing the inhabitants of the country was preceded either by a plural definite or by an indefinite article. It should be noted that in French (in which language the study was conducted), plural definites invite a generic reading (Robinson, 2005, p. 18), thereby favoring a generalized and taxonomic representation of the social category described. However, in the present experimental setting, plural definites remain referential, thus fulfilling the condition of a presupposition (i.e., referring to a salient referent in the context)[20]. As for indefinite NPs, they favor a single occurrence reading, thus presenting information about the social category as singular in the provided context[21].

[20] Robinson (2005, p. 18) points out that generic readings in French can be encoded either by singular or plural definite descriptions. One test allowing to claim the presence of a generic reading is to see if the predicate cannot apply to an individual (*Paul est rare [*Paul is rare]). In the present experimental setting, predicates can apply to an individual (e.g., Paul a joué au grand séducteur [Paul *played the great seducer*.]). This speaks in favor of a non-generic reading of the stimuli.

[21] Grice (1975, p. 56) provides examples with indefinite articles to illustrate the phenomenon of generalized conversational implicatures (e.g., "X is meeting a woman this evening."; "X went into a house yesterday and found a tortoise (…)"). He explains that the use of the indefinite article promotes the inference that the item is unfamiliar.

Following Sperber and Wilson (1986, p. 706), Experiment 2 tested whether readers would save processing efforts for presupposed contents, as opposed to asserted ones:

Hypothesis 4. Definite articles would be read more quickly than indefinite articles because they presuppose a referent which is highly salient (redundant in the context).

Finally, we conducted exploratory analyses to evaluate whether stereotype-incongruent information would be easier to process when introduced by an indefinite article (single occurrence reading) than by a definite one (generalized and definitional representation of the social category). These exploratory analyses aimed to evaluate whether stereotype-incongruent information was easier to process when it is under the scope of an indefinite description, as it promotes the reading of only one occurrence of an unexpected representation.

5.2. Experiment 1

Experiment 1 aimed to further assess the specific processing time course of gender stereotypes and to replicate previous findings (i.e., that gender stereotype-incongruent information is costly to process) cross-linguistically with French speaking Swiss participants (Hypothesis 1). Experiment 1 also investigated whether stereotype-incongruent information is costlier to process when it is presupposed through a definite description compared to when it is asserted though an indefinite description (Hypothesis 2).

5.2.1. Methods

5.2.1.1. Participants

For Experiment 1, 59 French speaking participants were recruited from a university in Switzerland. Only native French speakers were selected to participate in the experiment. The total sample size was set before data collection and based on the sample size estimation for "counterbalanced designs" developed by Westfall et al., (2014, p. 2026). The sample size estimation was conducted on Westfall et al.'s website (https://jakewestfall.shinyapps.io/). We used the "standard case" values of variance components (VPCs; Westfall et al., 2014, p. 2025), with a power set at .85, a medium effect size of $d = 0.50$, and a number of 22 stimuli. The sample size estimation revealed that 58.8 participants were required. No additional participant

In the present experimental design, the use of the indefinite article includes the notion of unfamiliarity. However, it also promotes a single occurrence reading.

was recruited once the pre-set sample size of 59 participants was reached. Following Müller and Mari (2021) and Singh et al. (2016), which employed the same experimental design as the current study, we excluded data from participants who had an accuracy rate for comprehension questions lower than 65%. This led to the exclusion of 2 participants. The final sample size resulted in 57 participants (31 women and 26 men; with an age mean of 23.87 years old, *SD* = 4.29).

5.2.1.2. Materials

The materials were constructed following a 2 x 2 design, manipulating (a) information about the social category, which either confirmed or violated stereotypical expectations, and (b) the NP introducing the social category, either with a definite article "*le/la*," "the" (presupposition condition), or with an indefinite article "*un/une*," "a/an" (assertion condition). The stimuli were created from the same model as those employed in Müller and Mari (2021) and Singh et al. (2016). Namely, the stimuli consisted in sets of two sentences written in French. The first sentence introduced a context, which was then followed by a target sentence matching or violating a gender stereotype. The target sentence introduced a specific agent marked grammatically by gender (e.g., chirurgien/chirurgienne, surgeon$_{male/female}$). The NP of the target sentence, i.e., the NP containing the social category concept, was introduced either with a definite article (working as a presupposition trigger) or an indefinite article (working as an assertion). In the end, each stimulus varied across four conditions which manipulated the effect of stereotypes and the article preceding the NP: (1) *stereotype-congruent and definite NP*, (2) *stereotype-congruent and indefinite NP*, (3) *stereotype-incongruent and definite NP*, and (4) *stereotype-incongruent and indefinite NP* (see Table 3).

Gender stereotypes were based on a selection of role nouns tested in Misersky et al. (2014) as well as additional role nouns commonly found in French speaking Switzerland. A list of 50 role nouns were pre-tested on another sample of 36 subjects (50% self-identified as women) from the same population as the participants of Experiment 1. The pre-test was run on Qualtrics (Provo, UT) and followed a procedure similar to Misersky et al. (2014). Participants had to indicate on a 5-point Likert scale their opinion about the extent to which role nouns consisted of women or men[22]. Response options included "mostly women," "more women," "as much women as men," "more men," "mostly men" (coded as 1 for "mostly women" and 5

[22] For example, *Veuillez indiquer si vous trouvez que plus de femmes ou d'hommes occupent la profession de chirurgiens/chirurgiennes,* [Please indicate whether you find that more women or men work as surgeons$_{male/female}$].

for "mostly men"). Role nouns that obtained the smallest scores (M = 2.23, SD = 0.32) were selected as female stereotypes and roles nouns that obtained highest scores (M = 3.79, SD = 0.38) were used as male stereotypes. In total, 22 stimuli were used, half related to female role nouns and half related to male role nouns. An additional set of 24 filler sentences was used to mask the purpose of the experiment. The complete list of stimuli and fillers is available at https://osf.io/b8h5q/.

Stimuli were also pre-tested in terms of plausibility. A total of 34 raters indicated, via Qualtrics (Provo, UT), the probability to encounter a specific social agent in a given situation (e.g., seeing surgeons in a hospital). The questions were asked in the following form: "*Si Marie va dans un hôpital, il est probable qu'elle rencontre … chirurgien.ne(s)*", "If Mary went to the hospital, it is likely that she encounters … surgeon(s)". Raters could choose between "zero," "only one," "one or more," "necessarily more than one" to replace the dots. For the selected stimuli, 76.4% of the raters chose "one or more"[23], assuring that the stimuli were considered as plausible.

5.2.1.3. Procedure

The experiment was created with and ran on E-Prime 2.0 software (Psychology Software Tools, Inc., 2012). We masked the purpose of the study from participants by informing them that they would participate in a study that investigated the links between causal information and its effects on the perception of narrativity in a reading task. Participants were instructed to read the sentences for comprehension. At the end of the study, the real purpose of the study was revealed.

Before running the experiment, participants were asked to indicate their age, gender, and mother tongue. The stimuli and fillers were then presented in sentence segments of 2-3 words (see Table 3), written in white 16-point Arial font on a black background. Each trial started with a white fixation cross on a black background, presented for 500 ms in the middle of the screen. The first segment then appeared on the screen. Participants would then press the spacebar to display the segments consecutively. This procedure prevented participants from displaying the whole sentence before reading it. Participants read only one condition of each stimulus, and as many stimuli from each of the four conditions, resulting in a within-subjects and within-stimuli design (Brauer & Curtin, 2018). Stimuli and fillers were presented randomly.

[23] 21.5% of the raters chose "necessarily more than one", 2.1% of the raters chose "only one," and none of the raters chose "zero."

Comprehension questions were used to assess whether participants remained attentive during the whole task. Comprehension questions were asked about the filler sentences only, and directly followed the corresponding filler. Participants answered yes or no by pressing on the "E" or "I" keys on the keyboard, according to the location of the yes/no answers on the screen. The experiment started with 6 practice trials, including one comprehension question, to familiarize participants with the task.

Table 3. Example of a stimulus of Experiment 1 in the four experimental conditions.

Condition	Context sentence	Target sentence
Stereotype-congruent and definite NP	Lucienne | est allée | à l'hôpital | le mois dernier. Lucienne | went to | the hospital | last month.	Le chirurgien | l'a opérée | avec une grande précision. The surgeon$_{male}$ | operated on her | with great precision.
Stereotype-congruent and indefinite NP	Lucienne | est allée | à l'hôpital | le mois dernier. Lucienne | went to | the hospital | last month.	Un chirurgien | l'a opérée | avec une grande précision. A surgeon$_{male}$ | operated on her | with great precision.
Stereotype-incongruent and definite NP	Lucienne | est allée | à l'hôpital | le mois dernier. Lucienne | went to | the hospital | last month.	La chirurgienne | l'a opérée | avec une grande précision. The surgeon$_{female}$ | operated on her | with great precision.
Stereotype-incongruent and indefinite NP	Lucienne | est allée | à l'hôpital | le mois dernier. Lucienne | went to | the hospital | last month.	Une chirurgienne | l'a opérée | avec une grande précision. A surgeon$_{female}$ | operated on her | with great precision.

Note. Vertical bars (|) indicate the separation between each segment.

5.2.2. Results

5.2.2.1. Data analysis

The effects of stereotype-congruent *versus* incongruent information and the article preceding the NP on information processing were measured by reading times, i.e., the time spent reading a sentence segment before clicking on the space bar to make a new segment appear. Three segments are considered for the analysis: (a) the critical segment consisting in the stereotype-congruent/incongruent information and the definite/indefinite NP, (b) the first spillover segment that follows the critical segment, and (c) the second spillover segment (see example (6); vertical bars separate the sentence segments). The two segments following the critical segment are traditionally included in the analysis of self-paced reading measures. In this way, it is possible to assess potential processing difficulties that emerged or persisted after reading the critical segment (Liversedge et al., 1998).

(6) Lucienne | est allée | à l'hôpital | le mois dernier. | La chirurgienne critical segment | l'a opérée spillover 1 | avec grande précision spillover 2.
[Lucienne | went to | the hospital | last month. | The surgeonfemale critical segment | operated on her spillover 1 | with great precision spillover 2.]

Reading times below 100 ms and above 4000 ms were excluded from the final dataset, leading to the suppression of 1.4% of data and a final dataset of 1238 datapoints (the dataset is available at https://osf.io/b8h5q/). The data were logarithmically transformed to meet the assumptions of mixed effects model analyses (i.e., homoscedasticity, linearity, and normality). Data analysis was conducted on RStudio (R Core Team, 2019, version 3.6.0), using the lme4 package (Bates, Mächler, et al., 2015).

5.2.2.2. Model selection

Model specification was driven by the experimental design, as recommended by experts in the field (Barr et al., 2013; Brauer & Curtin, 2018; Winter & Wieling, 2016). Fixed predictors are composed of the interaction between the stereotype condition (stereotype-congruent or incongruent information) and the NP condition (definite or indefinite article). Due to the repeated measures design, both subjects and stimuli created non-independence in the data and were thus included as by-subjects and by-stimuli random effects (Brauer & Curtin, 2018, p. 401). According to Barr et al. (2013), each fixed predictor that vary within-unit should include a random slope, as well as interactions when all factors vary within-units. In the present study, the stereotype condition and the NP condition varied both within-subjects and within-stimuli. Consequently, reading times were assessed with the following maximal mixed effect model: model <- lmer (log reading times ~ stereotype * NP + (stereotype + NP + stereotype*NP | subjects) + (stereotype + NP + stereotype*NP | stimuli)).

The maximal mixed effect model for the three analyzed segments converged. For the first spillover segment convergence was reached by using the built-in optimization procedure "bobyqa" of the lme4 package (Bates, Mächler, et al., 2015). This procedure has been acknowledged as one of the "remedies" that should be used to achieve convergence[24] (Brauer and Curtin 2018:404). The maximal mixed effect models for the three segments analyzed

[24] Failures of convergence are often due to the complexity of the random effect structure required by the experimental design. For the present study, the number of parameters estimates was 25, which might have been too high to reach a stable maximum likelihood estimation given the 1238 datapoints (Barr et al., 2013; Bates, Kliegl, et al., 2015; Brauer & Curtin, 2018; Winter, 2019).

resulted however in a singular fit. Singular fits are indicators that the models are overparametrized and that they should be reduced to parsimonious models, balancing at the same time the Type I error rate and statistical power (Bates, Kliegl, et al., 2015; Bates, Mächler, et al., 2015; Matuschek et al., 2017). We thus conducted a random effect Principal Component Analysis, using the rePCA function of the lme4 package (Bates, Mächler, et al., 2015). Goodness of fit was estimated with the likelihood ratio test (LRT) and AIC/BIC criteria (Bates, Kliegl, et al., 2015; Matuschek et al., 2017). The resulting models for reading times of the three segments are displayed in Table 4. The details of model selection and comparison are available at https://osf.io/b8h5q/. We also ran models including participants' gender to assess potential differences between self-identified male and female participants. For all three analyzed segments, we found no effect of gender. Gender was thus not included as a fixed predictor in the final models.

Table 4. Resulting parsimonious models for reading times on the three analyzed segments of Experiment 1.

Segment analyzed	Final parsimonious model
Critical segment	lmer(log critical segment ~ stereotype * NP + (stereotype*NP || subjects) + (stereotype || stimuli))
First spillover	lmer (log spillover1 ~ stereotype * NP + (stereotype + stereotype*NP || subjects) + (stereotype || stimuli))
Second spillover	lmer (log spillover2 ~ stereotype * NP + (stereotype + NP || subjects) + (stereotype || stimuli))

Note. Parsimonious models were selected after a random effect Principal Component Analysis, estimation of goodness of fit with likelihood ratio test, AIC, and BIC criteria (Bates, Kliegl, et al., 2015; Bates, Mächler, et al., 2015; Matuschek et al., 2017). Details of model selection are available at https://osf.io/b8h5q/.

5.2.2.3. Reading times for the critical segment

The effect of stereotype-congruent and incongruent information on reading times was first assessed. The analysis revealed that there was no main effect of stereotype on reading times of the critical segment. Although reading times of stereotype-incongruent information (M = 1106.44 ms, SD = 604.44) were longer than reading times of stereotype-congruent information (M = 1046.3 ms, SD = 571.2), this difference was not significant, $t(115.5) = -1.64$, $p = .103$ (see Table 6). When looking at the effect of definite and indefinite NPs only, we found no significant differences again between definite (M = 1067.1 ms, SD = 589.7) and indefinite NPs (M = 1085.5 ms, SD = 587.8), $t(1107) = 0.552$, $p = .581$ (see Table 6). No interaction effect

between stereotype information and the article preceding the NP were observed, $t(347.7) = 0.109$, $p = .913$ (see Table 6).

5.2.2.4. Reading times for the two spillover segments

The two segments following the critical segment were analyzed to assess whether a processing difficulty emerged after reading a particular segment (Liversedge et al., 1998).

Spillover 1. The analysis revealed a main effect of stereotype information on reading times for the first spillover segment, $t(82.12) = -2.4$, $p = .019$. Reading times of stereotype-incongruent information ($M = 852.61$ ms, $SD = 456.48$) were significantly longer than reading times of stereotype-congruent information ($M = 812.06$ ms, $SD = 420.55$). These results support Hypothesis 1, namely that information violating gender stereotypes is costlier to process than stereotype-congruent information. A main effect of the article was also observed on the first spillover, with longer reading times after definite NPs ($M = 848.11$ ms, $SD = 458.79$) than after indefinite NPs ($M = 816.66$ ms, $SD = 418.55$), $t(1076) = -2.18$, $p = .029$. No interaction effect between stereotype and NP was observed, $t(286.5) = 1.21$, $p = .229$ (see Table 6). Let us note that these results contradict the hypothesis of Relevance theory, namely that definite articles should be read more quickly than indefinite articles in plausible contexts. This issue is raised in the discussion section below. Contrast analyses were nonetheless conducted to assess whether within stereotype-incongruent conditions, longer processing times were observed with definite NPs as opposed to indefinite NPs (Hypothesis 2). These analyses revealed that reading times were significantly longer after reading stereotype-incongruent information introduced by a definite article ($M = 881.89$ ms, $SD = 480.11$) than when introduced by an indefinite article ($M = 824$ ms, $SD = 430.98$), $t(1076) = 2.18$, $p = .029$ (see Figure 1). Moreover, reading an incongruent stereotype introduced by a definite article was significantly costlier than in any other condition (see Table 5).

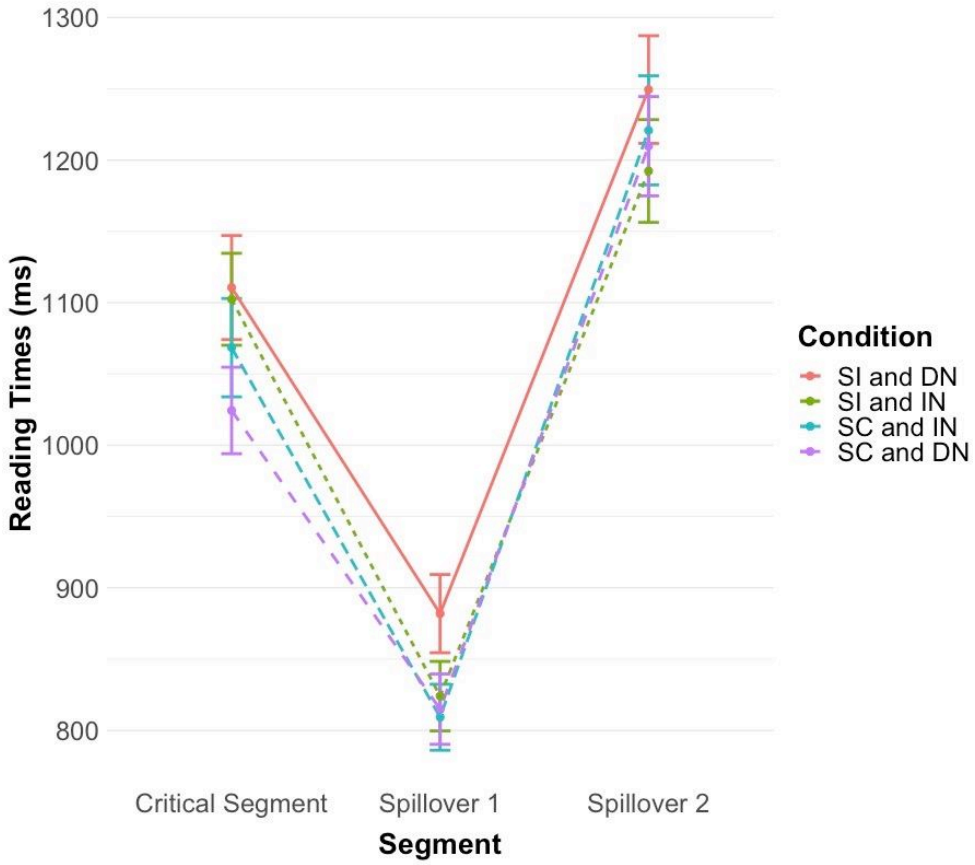
Spillover 2. The analysis of the second spillover segment revealed no effect of the stereotype information ($t(62.45) = -0.80$, $p = .429$), no effect of the NP ($t(178.65) = -1.33$, $p = .186$), and no interaction effect between stereotype and NP ($t(1048) = 0.79$, $p = .428$) (see Table 6). These results suggest that the difficulty of processing emerged right after reading the critical segment and stopped immediately after the first spillover, namely, once the verb phrase was reached (see Figure 4).

Table 5. Contrast analyses for the first spillover segment of Experiment 1

Conditions		**SC and IN**	**SC and DN**	**SI and IN**
	***M* (*SD*)**	***t*-test statistics**		
SC and IN	809 (*406*)			
SC and DN	815 (*435*)	$t(85.9) = 0.43$, $p = .672$		
SI and IN	824 (*431*)	$t(66.8) = 0.77$, $p = .447$	$t(79.2) = 0.38$, $p = .703$	
SI and DN	882 (*480*)	$t(66.7) = 2.72$, $p = .008$	$t(82.1) = 2.40$, $p = .019$	$t(1076) = 2.18$, $p = .029$

Note. SC = stereotype-congruent, SI = stereotype-incongruent; DN = definite NP, IN = indefinite NP.

Figure 4. Mean reading times in milliseconds for Experiment 1.



Note. Mean and standard error reading times (raw data) for each segment in Experiment 1. SC = stereotype-congruent, SI = stereotype-incongruent; DN = definite NP, IN = indefinite NP.

Table 6. Statistical results of the selected parsimonious models for Experiment 1.

Fixed effects							**Random effects**		
	Estimate	**SE**	**CI (95%)**	**t-value**	**DF**	**p-value**		**Var.**	**SD**
Critical segment									
(Intercept)	6.874	0.049	[6.78, 6.97]	142.94	89.82	<.001	Subjects intercept	0.09	0.30
Stereotype-congruent	-0.052	0.031	[-0.11, 0.11]	-1.64	115.5	.103	Subjects slope (stereotype*NP)	0.01	0.09
Indefinite NP	-0.017	0.030	[-0.04, 0.08]	0.55	1107	.581	Stimuli intercept	0.005	0.07
Interaction	-0.005	0.045	[-0.08, 0.09]	0.11	347.7	.913	Stimuli slope (stereotype)	0.001	0.04
First spillover									
(Intercept)	6.662	0.049	[6.56, 6.76]	137.03	76.99	<.001	Subjects intercept	0.11	0.33
Stereotype-congruent	-0.065	0.027	[-0.12, -0.01]	-2.40	82.12	**.019**	Subjects slope (stereotype)	0.004	0.06
Indefinite NP	-0.055	0.025	[-0.10, -0.005]	-2.18	1076	**.029**	Subjects slope (stereotype*NP)	0.003	0.05
Interaction	0.044	0.036	[-0.03, 0.12]	1.21	286.5	.229	Stimuli intercept	0.003	0.06
							Stimuli slope (stereotype)	0.001	0.03
Second spillover									
(Intercept)	7.008	0.062	[6.89, 7.13]	113.52	62.41	<.001	Subjects intercept	0.09	0.31
Stereotype-congruent	-0.026	0.033	[-0.09, 0.04]	-0.80	62.45	.429	Subjects slope (stereotype)	0.01	0.11
Indefinite NP	-0.039	0.029	[-0.09, 0.02]	-1.33	178.56	.186	Subjects slope (NP)	0.007	0.09
Interaction	0.030	0.038	[-0.04, 0.10]	0.79	1048	.428	Stimuli intercept	0.04	0.20
							Stimuli slope (stereotype)	0.004	0.06

Note. DF stands for degrees of freedom, SE for standard error, CI for confidence interval, Var. for Variance and SD for standard deviation. Values in bold are significant at $p < .05$ (calculated using Satterthwaites approximations). The selected mixed effects models are presented in Table 2.

5.2.3. Discussion

Experiment 1 replicated previous findings on the impact of gender stereotypes on processing times, using self-paced reading tasks. Unlike previous studies which analyzed reading times of complete sentences (e.g., Cacciari et al., 1997; Carreiras et al., 1996; Dickinson, 2011; Lassonde, 2015; Reynolds et al., 2006) or acceptability judgements (e.g., Gabriel et al., 2017; Garnham et al., 2002; Sato et al., 2013), this study presented sentence segments of 1-3 words, allowing a moment-by-moment analysis of processing difficulty. The analysis revealed that the processing difficulty of stereotype-incongruent role nouns was delayed to the first spillover segment. This is in line with eye-tracking studies showing that reading times are significantly slowed down upon and/or after reading a pronoun that led to a mismatch between a role noun and its anaphoric pronoun (e.g., reading "electrician" followed by "she," Reali et al. 2015, see also Duffy & Keir, 2004; Esaulova et al., 2014; Irmen, 2007; Kennison & Trofe, 2003). Experiment 1 thus replicates previous findings with French speaking Swiss participants: Information violating gender stereotypes is costlier to process than stereotype-congruent information.

Turning now to the effect of NPs, Experiment 1 revealed that definite NPs led to longer reading times than indefinite NPs on the spillover region. As noted above, these results contradict the assumption of Relevance theory, according to which definite articles should be less costly to process than indefinite articles within a plausible context. However, it should be noted that the observed longer processing time of definite articles was mainly driven by the processing of stereotype-incongruent information which generated a significant slowdown. Indeed, as revealed by contrast analyses, stereotype-incongruent information introduced by a definite NP (e.g., the surgeon$_{female}$) were significantly costlier than all other conditions. Within stereotype-congruent conditions (e.g., surgeon$_{male}$) definite articles were slightly (6 ms) costlier to process than indefinite articles. This is in line with Singh et al. (2016, p. 617), who also observed a slight slowdown with definite articles, as opposed to indefinite ones. It is likely that this experimental setup makes the processing of the definite articles costly, due to a difficulty to identify the referent in the previous context sentence. As we pointed out above (Section 5.1.2), the stimuli required a bridging inference, which is not necessary when the noun is preceded by an indefinite article, as it merely introduces a new referent.

Together these findings show that the processing of definite NPs, which requires the identification of a salient referent, is significantly affected by stereotypical representations.

Normally, in plausible contexts, definite NPs should require little processing efforts (Sperber & Wilson, 1986, p. 706). However, the present experimental design suggests that the processing of definite descriptions interacts with social cognitive modules, generating a significant slowdown, despite a plausible context.

In sum, Hypothesis 1 was confirmed: Stereotypes are predictive of linguistic processing, where information incongruent with gender stereotypes is significantly costlier to process than stereotype-congruent information. Furthermore, Hypothesis 2 was also confirmed: Definite NPs were significantly costlier than indefinite NPs within the incongruent-stereotype condition. Regarding Hypothesis 2, it should be stressed that previous experiments on context plausibility (Singh et al. 2016; Müller and Mari 2021) were not able to show a significant difference between definite and indefinite articles within implausible condition.

5.3. Experiment 2

Experiment 2 assessed (a) whether information violating expectations about secondary social categories (nationality stereotypes) are costly to process as is information violating gender stereotypes (Hypothesis 3), and (b) whether definite articles are more quickly read than indefinite articles in redundant contexts (Hypothesis 4). Exploratory analyses were conducted to evaluate whether stereotype-incongruent information is easier to process when introduced by an indefinite article (single occurrence reading) than by a definite one (generalized representation of the social category).

5.3.1. Methods

5.3.1.1. Participants

For Experiment 2, 49 French speaking participants were recruited from a university in Switzerland. As in Experiment 1, only native French speakers were selected to participate in the experiment. The total sample size was set before data collection and based on a sample size estimation as conducted for Experiment 1. Using the website of Westfall et al. (2014; https://jakewestfall.shinyapps.io/), we set the values for "counterbalanced designs" as in the two previous experiments, namely with the "standard case" values of VPCs, a power of .80, a medium effect size of $d = 0.50$ and a number of stimuli of 20. The sample size estimation revealed that 48.8 participants were required. No additional participant was recruited once the

pre-set sample size of 49 participants was reached. Similar to Experiment 1, we controlled that participants provided a minimum of 65% accuracy rate for comprehension questions. All participants responded with more than 65% accuracy. The final sample size resulted in 49 participants (28 women and 21 men; with an age mean of 23.06 years old, *SD* = 3.53).

5.3.1.2. Materials

The stimuli were constructed in a similar way to Experiment 1. They consisted in two sentences written in French, with the first sentence introducing the context, and the following sentence matching or violating a nationality-related stereotypes. The target sentence introduced a social category, i.e., inhabitants of a country. The NP introducing the social category, was either a definite NP (working as a presupposition, favoring a generalized and taxonomic representation of the social category) or an indefinite NP (working as an assertion, favoring a single occurrence reading of the stereotype). As in Experiment 1, each stimulus varied across four conditions: (1) *stereotype-congruent with definite NP*, (2) *stereotype-congruent with indefinite NP*, (3) *stereotype-incongruent with definite NP*, and (4) *stereotype-incongruent with indefinite NP* (see Table 7).

Nationality-related stereotypes were based on folk stereotypes found in everyday Sspeech (e.g., in movies, jokes, hearsay, comics, etc.) in the region of French speaking Switzerland. A list of 90 nationality stereotypes were pre-tested on another sample of 36 subjects (50% self-identified as women) from the same population as the final sample of Experiment 2. The pre-test was run on Qualtrics (Provo, UT) and asked participants to indicate on a 5-point Likert scale their opinion about diverse statements[25]. Response options included "agree," "somewhat agree," "neither agree nor disagree," "somewhat disagree," and "disagree" (coded as 1 for "agree" and 5 for "disagree"). Statements that obtained the smallest scores (*M* = 2.35, *SD* = 0.33) were selected as nationality stereotypes and statements that obtained the highest scores (*M* = 3.87, *SD* = 0.29) were used as nationality counter-stereotype in the present study. In total, we used 20 stimuli, half matching nationality stereotype and half violating nationality stereotypes[26]. An additional set of 24 filler sentences was used to veil the purpose of the experiment. The complete list of stimuli and fillers is available at https://osf.io/b8h5q/.

[25] For example, *À quel point êtes-vous d'accord avec la proposition suivante: "les Japonais sont de grands séducteurs"* [To what extent do you agree with the following statement "the Japanese are great seducers"].
[26] The plausibility was not pre-tested for Experiment 2, because the inhabitants introduced in the target sentence corresponded to those of the country presented in the context sentence (e.g., going to Japan and seeing Japanese is highly plausible).

Table 7. Example of a stimulus of Experiment 2 in the four experimental conditions.

Condition	Context sentence	Target sentence
Stereotype-congruent and definite NP	Mathilde | est allée | en Italie | le week-end dernier. Mathilde | went to | Italy | last weekend.	Les Italiens | ont joué | aux grands séducteurs | durant tout le séjour. The Italians | played | the great seducers | during the whole stay.
Stereotype-congruent and indefinite NP	Mathilde | est allée | en Italie | le week-end dernier. Mathilde | went to | Italy | last weekend.	Un Italien | a joué | au grand séducteur | durant tout le séjour. An Italian | played | the great seducer | during the whole stay.
Stereotype-incongruent and definite NP	Mathilde | est allée | au Japon | le week-end dernier. Mathilde | went to | Japan | last weekend.	Les Japonais | ont joué | aux grands séducteurs | durant tout le séjour. The Japanese | played | the great seducers | during the whole stay.
Stereotype-incongruent and indefinite NP	Mathilde | est allée | au Japon | le week-end dernier. Mathilde | went to | Japan | last weekend.	Un Japonais | a joué | au grand séducteur | durant tout le séjour. A Japanese | played | the great seducer | during the whole stay.

Note. Vertical bars (\|) mark presentation boundaries (i.e., sentences segments).

5.3.1.3. Procedure

The procedure was the same as the one described in Experiment 1.

5.3.2. Results

5.3.2.1. Data analysis

As in Experiment 1, the effects of stereotype-congruent *versus* incongruent information and the definite *versus* indefinite article were measured by reading times. Four segments were considered for the analysis: (a) the one containing the definite/indefinite NP that introduced the social category (i.e., inhabitants of a country), (b) the spillover segment to assess potential persistence of processing difficulty, (c) the segment presenting stereotype-congruent/incongruent information, and (d) its spillover segment (see example (7); vertical bars separate the sentence segments):

(7) Mathilde | est allée | au Japon | le week-end dernier. | Les Japonais$_{\text{critical segment 1}}$ | ont joué$_{\text{spillover 1}}$ | aux grands séducteurs$_{\text{critical segment 2}}$ | durant tout le séjour$_{\text{spillover 2}}$.

[Mathilde | went to | Japan | last weekend. | The Japanse$_{\text{critical segment 1}}$ | played$_{\text{spillover 1}}$ | the great seducers$_{\text{critical segment 2}}$ | during the whole stay$_{\text{spillover 2}}$.]

Similar to Experiment 1, reading times below 100 ms and above 4000 ms have been excluded from the final dataset. This data exclusion resulted in the suppression of 1.1% of data and a final dataset of 969 datapoints (the dataset is available at https://osf.io/b8h5q/). The data were logarithmically transformed to meet the assumptions of mixed effects model analyses and data analysis was conducted on Rstudio (R Core Team, 2019, version 3.6.0), using the lme4 package (Bates, Mächler, et al., 2015).

5.3.2.2. Model selection

We followed the same procedure as in Experiment 1 to specify the model (i.e., model selection based on the experimental design). The first segment under investigation in the present experiment did not mix the types of NPs and stereotype information. As illustrated in example (7), the first segment varies only in terms of the article used, namely definite or indefinite. The information violating / confirming stereotypes is only introduced in the seventh segment (critical segment 2). As a consequence, reading times on the first critical segment and the first spillover were assessed with the following maximal mixed effect model: model1 <- lmer (log reading times ~ NP + (NP | subjects) + (NP | stimuli)). On the other hand, reading times of the second critical segment and its spillover could be affected by both the type of NP and stereotype information. Therefore, reading times of those remaining segments were analyzed with the following maximal mixed effect model: model2 <- lmer (log reading times ~ stereotype * NP + (stereotype*NP | subjects) + (stereotype*NP | stimuli)).

The maximal mixed effect model for the four analyzed segments reached convergence. For the two critical segments and the two spillover segments, the built-in optimization procedures "nlminbwrap" and "bobyqa" of the lme4 package (Bates, Mächler, et al., 2015) were used, respectively. The maximal mixed effect models for the four segments resulted however in a singular fit, indicating that the models were overparametrized. We thus conducted a random effect Principal Component Analysis, using the rePCA function of the lme4 package (Bates, Mächler, et al., 2015). Goodness of fit was estimated with the likelihood ratio test (LRT) and AIC/BIC criteria (Bates, Kliegl, et al., 2015; Matuschek et al., 2017). The resulting models for reading times of the four segments are displayed in Table 8. The details of model selection and comparison are available at https://osf.io/b8h5q/.

Table 8. Resulting parsimonious models for reading times on the three analyzed segments of Experiment 2.

Segment analyzed	Final parsimonious model
Critical segment 1	lmer (log critical segment 1 ~ NP + (NP || subjects) + (0 + NP || stimuli))
Spillover 1	lmer (log spillover1 ~ NP + (NP || subjects) + (1 | stimuli))
Critical segment 2	lmer(log critical segment 2 ~ stereotype * NP + (1 | subjects) + (stereotype + NP || stimuli))
Spillover 2	lmer(log spillover 2 ~ stereotype * NP + (NP || subjects) + (NP || stimuli))

Note. Parsimonious models were selected after a random effect Principal Component Analysis, estimation of goodness of fit with likelihood ratio test, AIC, and BIC criteria (Bates, Kliegl, et al., 2015; Bates, Mächler, et al., 2015; Matuschek et al., 2017). Details of model selection are available at https://osf.io/b8h5q/.

5.3.2.3. Reading times for the first critical and spillover segments

The analysis revealed that there was no main effect of definite or indefinite NPs on reading times of the critical segment, $t(44.23) = 1.59$, $p = .118$. The next segment was also analyzed to assess if a processing difficulty emerged after reading the definite/indefinite NPs. The analyses revealed a significant difference between reading times, $t(51.24) = 2.49$, $p = .016$, with longer reading times following an indefinite NP ($M = 884.69$ ms, $SD = 493.92$) compared to a definite NP ($M = 825.08$ ms, $SD = 439.37$). This result confirms Hypothesis 4, where definite NPs were expected to be read more quickly than indefinite ones (see Table 9). These results are compatible with Relevance theory, which argues that definite descriptions allow to spare cognitive efforts (compared with indefinite descriptions).

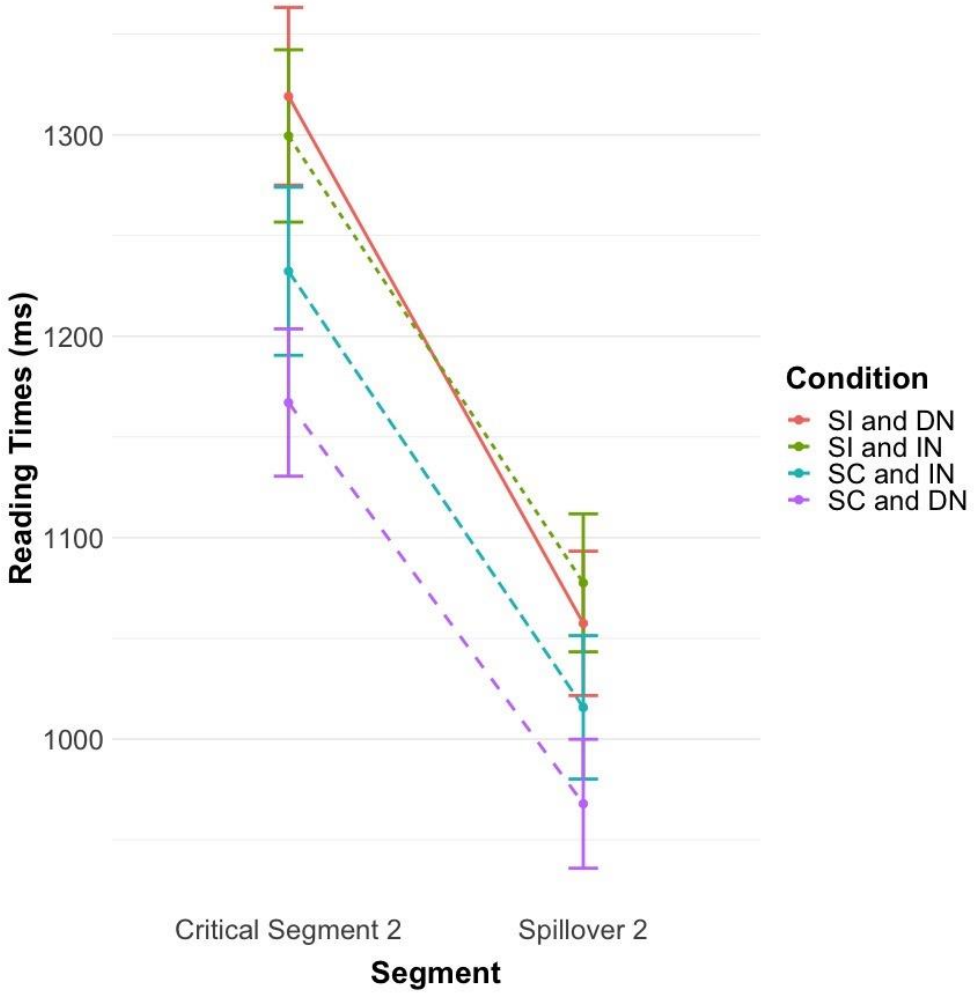
5.3.2.4. Reading times for the second critical and spillover segments

The analysis revealed that there was a main effect of stereotype information on reading times of the second critical segment. Reading times of stereotype-incongruent information ($M = 1309.26$ ms, $SD = 676.07$) were significantly longer than reading times of stereotype-congruent information ($M = 1199.43$ ms, $SD = 610.02$), $t(51.6) = -3.05$, $p = .004$. A main effect of stereotype was also observed for the spillover segment, with longer reading times for the stereotype-incongruent condition ($M = 1066.05$ ms, $SD = 538.86$) compared to the stereotype-congruent condition ($M = 993.24$ ms, $SD = 523.52$), $t(831.7) = -3.04$, $p = .002$ (see Figure 5). These results support Hypothesis 3, according to which stereotype-incongruent information elicit longer reading times than stereotype-congruent information. Importantly, these results provide evidence for the persistence of stereotype effects with secondary social categories. No

effect of definite/indefinite NPs and no interaction effect were observed for both the second critical segment and its spillover (see Table 9).

Contrast analyses were nonetheless conducted on both segments to explore the possibility that stereotype-incongruent information might be easier to process when introduced by an indefinite article as opposed to a definite article. These analyses revealed that reading times in the *stereotype-incongruent with indefinite NP* condition were not significantly faster than in the *stereotype-incongruent with definite NP* condition, $t(64.7) = 0.27$, $p = .786$ (critical segment 2) and $t(49.7) = 0.43$, $p = .666$ (second spillover). In other words, the processing of stereotype-incongruent information does not appear to be affected by linguistic markers of definiteness (see Figure 5).

Figure 5. Mean reading times in milliseconds for Experiment 2 (second critical segment and spillover).



Note. Mean and standard error reading times (raw data) in Experiment 2 for the second critical segment and its spillover. SC = stereotype-congruent, SI = stereotype-incongruent; DN = definite NP, IN = indefinite NP.

Table 9. Statistical results of the selected parsimonious models for Experiment 2.

Fixed effects							**Random effects**		
	Estimate	**SE**	**CI (95%)**	**t-value**	**DF**	**p-value**		**Var.**	**SD**
Critical segment 1									
(Intercept)	6.783	0.044	[6.69, 6.87]	152.78	52.98	<.001	Subjects intercept	0.08	0.29
Indefinite NP	0.050	0.031	[-0.01, 0.11]	1.59	44.23	.118	Subjects slope (NP)	0.02	0.13
							Stimuli (NP)	0.002	0.05
Spillover 1									
(Intercept)	6.600	0.051	[6.50, 6.70]	129.10	55.97	<.001	Subjects intercept	0.11	0.33
Indefinite NP	0.060	0.024	[0.01, 0.11]	2.49	51.24	**.016**	Subjects slope (NP)	0.006	0.07
							Stimuli intercept	0.002	0.05
Critical segment 2									
(Intercept)	7.063	0.059	[6.95, 7.18]	119.07	71.81	<.001	Subjects intercept	0.10	0.32
Stereotype-congruent	-0.103	0.034	[-0.17, -0.03]	-3.05	51.61	**.004**	Stimuli intercept	0.02	0.14
Indefinite NP	-0.008	0.031	[-0.07, 0.05]	-0.27	69.63	.787	Stimuli slope (stereotype)	0.005	0.07
Interaction	0.053	0.042	[-0.03, 0.14]	1.26	864.35	.207	Stimuli slope (NP)	0.002	0.04
Spillover 2									
(Intercept)	6.864	0.051	[6.76, 6.97]	133.79	70.98	<.001	Subjects intercept	0.09	0.31
Stereotype-congruent	-0.085	0.028	[-0.14, -0.03]	-3.04	831.65	**.002**	Subjects slope (NP)	0.008	0.09
Indefinite NP	0.015	0.033	[-0.05, 0.08]	0.45	67.86	.654	Stimuli intercept	0.006	0.08
Interaction	0.026	0.039	[-0.05, 0.10]	0.648	831.09	.517	Stimuli slope (NP)	0.09	0.31

Note. DF stands for degrees of freedom, SE for standard error, CI for confidence interval, Var. for Variance and SD for standard deviation. Values in bold are significant at $p < .05$ (calculated using Satterthwaites approximations). The selected mixed effects models are presented in Table 6.

5.3.3. Discussion

Experiment 2 revealed that stereotype-incongruent information about nationalities is longer to process than stereotype-congruent information. This finding provides further evidence that information confirming stereotypical expectations is easily processed, whereas information violating stereotypical expectations about secondary social categories is difficult to process (Hypothesis 3). Interestingly, the effects of information confirming / violating nationality stereotype already appeared on the critical segment and persisted in the spillover segment.

Regarding the effects of definite/indefinite NPs, Experiment 2 tested whether indefinite descriptions were costlier to process than definite descriptions, as proposed by Relevance theory. The present study confirmed Hypothesis 4, showing that definite NPs led to faster reading times than indefinite ones. This finding is all the more interesting in light of Experiment 1, where the fast reading of definite descriptions was disrupted because of information violating gender stereotypes. Finally, we explored whether indefinite articles (i.e., the representation of a single occurrence within a kind) could facilitate the processing of stereotype-incongruent information. These analyses revealed that indefinite articles could not make stereotype-incongruent information easier to process than when subjected to a generalization (i.e., a plural definite). Thus, in this experimental setup, the processing effects of stereotypes appear to be stronger than linguistic markers.

5.4. General discussion

The present study investigated the effects of social modules of cognition on the relevance-guided comprehension heuristic across two experiments, in order to shed light on the relevance comprehension heuristic. Both experiments assessed the extent to which stereotypes impact the processing of specific linguistic information. Experiment 1 aimed to replicate previous findings on the effects of gender stereotypes on reading cross-linguistically. Experiment 2 sought to investigate the effects of secondary social categories, i.e., nationalities.

The results of Experiment 1 showed that information violating gender stereotypes is longer to process than stereotype-congruent information (Hypothesis 1). This finding goes in line with previous studies that investigated, in different languages, the effect of gender stereotypes on reading and anaphora resolution (cf. Section 5.1.1). Our findings with French speaking Swiss participants support the cross-linguistic evidence that gender is rapidly encoded

during reading (Garnham et al., 2002; Gygax et al., 2021) and affects processing depending on whether the information communicated matches one's stereotypical expectations.

Furthermore, Experiment 1 revealed that stereotype-incongruent information makes the processing of presuppositional contents (definite articles) significantly costlier than assertions (indefinite descriptions) (Hypothesis 2). This is because the identification of a salient referent (required for definite NPs) is inconsistent with the encoding of stereotype-incongruent information. Importantly, these findings offer promising opportunities for the study of the relevance comprehension heuristic. While previous studies failed to reach conclusive results when using a broad category of plausibility (Müller & Mari, 2021; Singh et al., 2016), the present study revealed that the processing of definite descriptions is affected by stereotype information, generating a significant slowdown in narrowly defined plausible contexts. We suggest that the "plausibility of contexts," used in these previous studies, conflated different variables, such as surprise effects, comprehension problems as well as typicality effects.

In Experiment 2, we further assessed the effect of stereotypes about secondary social categories on processing, and revealed that information violating nationality stereotypes was costly to process (Hypothesis 3). This is consistent with Lassonde's (2015) study, showing that sentences containing stereotype-incongruent information about diverse social categories (e.g., nuns, rockstars) are costly to process. Together, these findings have some interesting implications, suggesting that any kind of stereotype-incongruent information would be difficult to process because more cognitive effort is required to access the information from associative memory. Unfortunately, this possibility has rarely been addressed, as most studies to date have focused only on gender stereotypes. In an endeavor to determine whether the effects of stereotypes on processing are consistent, future studies should investigate, with varying methodologies (e.g., response times, self-paced reading, eye-tracking, or event-related brain potentials) and across cultures, how and whether information processing is similarly affected by stereotype about various social categories.

Experiment 2 also revealed that definite descriptions are less costly than indefinite descriptions when the context is redundant. These results align with Relevance theory, which argues that definite descriptions allow to spare cognitive efforts (compared with indefinite descriptions). Moreover, our exploratory analysis showed that when a single occurrence was encoded linguistically (e.g., "A Japanese played the great seducer", as opposed to "The Japanese played the great seducers"), it did not facilitate the processing of incongruent-

stereotype information. This suggests that the processing of incongruent-stereotype information cannot be modulated by linguistic markers[27].

Before concluding, let us note that the present study's findings bear some important considerations. Given the current context and issues, it seems particularly important to study how stereotypical information is processed. For instance, although much effort and attention has been paid to gender equality in the 21st century (e.g., the increasing use of inclusive language, the promotion of STEM professions among girls and women, or strikes for women's right), we still observed that some conceptions of gender roles remain unchanged. Moreover, current crises (namely the COVID-19 pandemic, the war in Ukraine, or the climate crisis) led people to rapidly form negative stereotypes about inhabitants of certain countries. By documenting how people process stereotypes during reading, the present study showed that information confirming a stereotype is easily processed and thus, might not be questioned or noticed, while still being significant in the relevance comprehension heuristic. Furthermore, this could play a role in the maintenance of stereotypical expectations and the emergence of prejudices.

We also stress two important limitations to the current study. First, it does not allow to make direct comparisons between the two tested stereotypes. Indeed, gender stereotype in/congruence occurred within the NP (e.g., *"*une chirurgienne*"; "*a surgeon$_{female}$*"*) whereas nationality stereotype in/congruence stood in the relation between the NP and the predicate (e.g., "A/The Japanese *played the great seducer/s"*). In a future study, it would be worth testing these two stereotypes, and others, in a comparable way. Furthermore, our study focused only on narrow linguistic phenomena (definite *versus* indefinite descriptions). While this may be an asset experimentally (limiting other variables weighing in the processing speed), further studies are needed to see if stereotypes also constrain the processing of other linguistic markers, such as other presupposition triggers.

Overall, the present study's findings suggest that stereotypes bring significant constraints on the processing of linguistic information. These elements are of interest for Relevance theory, insofar as they confirm that the comprehension heuristic is constrained by information which goes beyond propositional cues, such as the listener's knowledge about

[27] Let us note that an anonymous reviewer drew our attention to the potential problems of confounding variables in the experimental setup. For this reason, further experiments with a setup that better isolates the variables should be conducted to test this hypothesis.

social categories. Across both experiments, stereotype-incongruent information was less salient than stereotype-congruent ones, making it cognitively more costly to process. Moreover, these findings suggest a possible hierarchy between social and linguistic information in the derivation of meaning: Indeed, while stereotype-incongruent information slows down the processing of definite descriptions (which are normally processed quickly), we did not observe that the processing of stereotype-incongruent information was facilitated when preceded by an indefinite article (single occurrence reading).

Summary and comments

Chapter 4 presented the results of a collaborative research, bridging the fields of social cognition and pragmatics. Drawing from theories on stereotype processing and Relevance theory, two experiments assessed how specific inputs affected information processing during reading. The findings revealed that stereotype-incongruent information about gender and nationality categories were costly to process, and that this observed processing difficulty was exacerbated when stereotype-incongruent information was presupposed by definite descriptions.

Furthermore, we analyzed whether stereotype-incongruent information was easier to process when introduced by a singular indefinite article (which promotes single occurrence interpretation) than by a plural definite one (which promotes generalized and definitional interpretation). The analysis revealed that there was no significant difference in reading times between the two kinds of articles. This suggests that readers encountered similar processing difficulties for stereotype-incongruent items when the article promoted the representation of a single occurrence within a kind (i.e., the retrieval of a particular exemplar) as when it promoted a generalized representation of a social category. Unfortunately, these effects could only be assessed within Experiment 2 that manipulated both single occurrence and generalized interpretations. Direct comparisons of the effects of single occurrence *versus* generalized interpretations on reading times for gender stereo types was not possible in Experiment 1 as it only employed singular articles.

Moreover, because this research was conducted in the context of an inter-disciplinary project, Chapter 4 did not provide a complete and precise account of how stereotype activation impacts information processing. Chapter 5 attempts to overcome this lack by presenting a more detailed, moment-by-moment assessment of stereotype processing during reading. Unlike

Chapter 4, the next chapter focuses mainly on the effects of stereotypes on information processing and puts less emphasis on the effects of presupposition.

6. Chapter 5

The effects of stereotypes on information processing: An assessment with eye-tracking methodology[28]

Abstract

How information processing is affected by stereotypical expectations during reading is not yet fully understood. While past research demonstrated that stereotype-incongruent information elicits longer reading times, these earlier works failed to provide a rigorous assessment of the specific processing time course of stereotype information and how processing can be modulated by specific linguistic markers. Cognitive approaches to stereotyping postulate that reading about category members automatically activates stereotypical expectancies about the likely attributes of this social category. Theories in pragmatics posit that presupposed contents, introduced with a definite article, are costly to process when the identification of a salient referent is inconsistent with the encoding of contextual information, i.e., when encountering stereotype-incongruent information. The present study presents two experiments that test those two hypotheses, by assessing with an eye-tracker how reading of French speaking Swiss adults (Experiment 1, N = 54; Experiment 2, N = 45) was affected by stereotype-incongruent information and presupposed contents. Results showed that participants slowed down and interrupted their forward reading of the text after encountering *gender* stereotype-incongruent information. These effects were only marginally significant when information violated *nationality*-related stereotypes. Overall, the present study showed that when information violated stereotypical expectations about gender roles, the forward progress of reading was interrupted, and attention shifted back to the locus of processing difficulty.

6.1. Introduction

This chapter presents a further assessment of the specific processing time course of stereotype-incongruent information and presupposed contents by testing two theoretical

[28] Unpublished research conducted in collaboration with Misha-Laura Müller, in partial fulfillment of our respective Doctorate degrees.

perspectives from social psychology and pragmatics. The rationale for the present study was the same as the one presented in Chapter 4. However, while the previous study used self-paced reading tasks, the present study employed an eye-tracking methodology to provide a finer moment-by-moment analysis of processing in a natural reading setting. Moreover, the emphasis in this chapter is on the effects of stereotypes on information processing, rather than on the effects of definiteness. Theoretical accounts in social psychology postulate that stereotypes get automatically activated from semantic memory in the presence of a category member (e.g., when reading about some social categories) (Gilbert & Hixon, 1991; Krieglmeyer & Sherman, 2012; Quadflieg & Macrae, 2011). This assumption rests on associative networks models (Anderson & Bower, 1973; Collins & Loftus, 1975; Wyer & Carlston, 2018), insofar as the activation of a concept (e.g., Swiss) travels through associated links to activate related concepts (e.g., punctual). When information contradicts a stereotype, readers will need more time to integrate this unexpected item into their mental representation of the text (because the activated concepts are not related), resulting in increased processing difficulty (Rees et al., 2020). As long as the activated concepts are related, information processing operates automatically; when information is unexpected, ongoing processing is interrupted and attention is likely to shift to the unexpected item (Reisenzein et al., 2019). Similarly, theories in pragmatics posit that presupposed contents are costly to process when the identification of a salient referent is impossible in a given the context, as in example (1); in contrast, processing is facilitated when the presupposed contents occur in a compatible context, as in example (2) (Singh et al., 2016; Sperber & Wilson, 1986).

(1) Mary went to the office a few hours ago. The lifeguard warned her there about the weather.
(2) Mary went to the beach a few hours ago. The lifeguard warned her there about the weather.

(Singh et al. 2016, p. 631)

The previous study, using self-paced reading tasks, revealed that (a) stereotype-incongruent information was indeed costly to process, (b) presupposed contents were easily processed when the context was compatible, and (c) processing stereotype-incongruent information was significantly costlier when it was presupposed compared to asserted (e.g., processing "*The* surgeon$_{female}$ operated on her" was costlier than processing "*A* surgeon$_{female}$ operated on her").

The present study aimed to replicate and specify these previously observed effects with eye-tracking methodology.

Eye-tracking methodology benefits from some valuable advantages and can provide finer insights on information processing that cannot be revealed with other methods. First, eye-trackers record a person's gaze in real time, which enables to detect, moment-by-moment, the locus of attention as well as the time to process a word or series of words (Eckstein et al., 2017; Liversedge et al., 1998). Second, eye-tracking methods record readers' gaze in almost natural settings – while reading on a computer screen – which allows to measure information processing without additional costs due to task demands (Rayner, 1998). Importantly, reading pace can be greatly influenced by the type of experimental task employed. For instance, it has been shown that reading pace in self-paced reading tasks was half less rapid than during normal reading (Rayner, 1998). Finally, eye-tracking methods allow to distinguish online from offline processing depending on the time course of reading. On the one hand, online processing corresponds to the immediate processing of a word (or series of words), when a text region is read for the first time. As such, online processing measures are "sensitive to processing difficulty experienced immediately on reading [a] word" (Liversedge et al., 1998a, p. 58). On the other hand, offline processing corresponds to late processing of a text region after having been read for the first time and provides evidence for a delayed effect of the manipulation (Liversedge et al., 1998).

The next two sections present key principles in eye movement research and the measures traditionally used to assess information processing during reading. Section 6.1.3 integrates theoretical perspectives on stereotyping and eye movement control. Section 6.1.4 reviews research that used eye movements to assess stereotype activation and presents a related line of research focusing on semantic violation in reading. The remaining sections present two experimental studies and discuss their respective findings.

6.1.1. Key principles of eye movement research

A fundamental principle in eye movement research is that the duration of fixations coincides with the cognitive processing involved when reading a specific text region (Inhoff & Radach, 1998). The duration of fixations will increase as a text region becomes more difficult to process, be it because of frequency effects (words occurring less frequently in a language), predictability effects (low predictability of words in a given context) or other word properties

(e.g., length, orthography) (Kliegl et al., 2012). Different eye movement measures are used to analyze fixation durations, and thus to assess information processing. As previously presented, eye movement measures can either reflect online or offline processing. Measures of online processing include first fixation duration, first-pass duration, and regression path time. To assess offline processing, total gaze duration and total sentence processing are commonly computed. These measures are defined in Table 10.

Table 10. Eye movement measures traditionally used to assess information processing during reading.

Eye movement measure	Definition
Online processing measures	
First fixation duration	The duration of the first fixation occurring on a word or text region.
First-pass duration	The sum of all fixations on a word or text region from the moment when the gaze entered the region to the left until the gaze left the region in any direction.
Regression path time	The sum of all fixation durations on a word or text region from entering it on the left and leaving it to the right, including any fixation made on previous regions.
Offline processing measures	
Total gaze duration	The sum of all fixation durations occurring on a word or text region while the stimulus was displayed.
Total sentence processing	The total duration for reading the sentence containing the critical word or text region.

First fixation and first-pass durations are indicators of very early processing (Warren, 2011). Regression path time is an indicator of comprehension failures. Specifically, when readers encounter a text region that contradicts their prior interpretation of the sentence, they make a regression as soon as they read disambiguating information (Blanchard & Iran-Nejad, 1987; Frazier & Rayner, 1982; Rayner, 1998; Vauras et al., 1992). Total gaze duration and sentence processing are sensitive to processes that occur after a text region has been read and reflect the total amount of processing that a text region or sentence receives (Daneman et al., 1995; Warren, 2011).

6.1.2. Common processes involved in eye movement control and stereotyping

To comprehend a text, readers combine elements that are derived explicitly from the text, as well as elements that are implicit, coming from their previously acquired knowledge (Elbro & Buch-Iversen, 2013; Graesser et al., 1994; Gygax et al., 2021; Kendeou et al., 2016). Initial theories of eye movement control during reading only accounted for the effect of lexical

processing but ignored the effects of higher-level language processing (Warren, 2011). The E-Z Reader model of eye movement control in reading (Reichle, 2011; Reichle et al., 1998, 2003, 2009) provides an account that integrates perceptual, cognitive, and motor processes. According to the E-Z Reader model, the eyes move forward when the currently attended word has been processed sufficiently to move on to the next one. The key assumption is that higher-order processes (e.g., integrating new information) lag behind and interrupt the forward progress of the eyes when information violates the readers' knowledge or expectations (Juhasz & Pollastek, 2011; Reichle et al., 2009; Warren, 2011). In brief, the E-Z Reader model proposes that when the fit between a text region and a context is poor, e.g., when information violates readers' expectations, forward saccades are cancelled. Instead, fixations remain on the critical region and a regression is programmed (Reichle, 2011; Reichle et al., 1998, 2003, 2009). Crucially, the E-Z Reader model predicts that forward eye movements will be suppressed, and regressions will be initiated when information violates the readers' expectations.

Similarly, theoretical accounts of stereotyping (Sherman et al., 1998; 2013) propose that information consistent with a stereotype is processed effortlessly, in "a conceptually fluent" way. Stereotype-congruent information is easily processed and well comprehended, even under limited cognitive resources. In this way, perceivers do not waste attentional resources on familiar expected items. However, in the presence of unexpected information (i.e., stereotype-incongruent information), residual attentional resources are redirected to enable perceivers to process this potentially meaningful piece of information (Sherman et al., 1998; see also Macrae & Bodenhausen, 2000; Reisenzein et al., 2019). Importantly, this account builds on a cognitive approach of stereotyping, positing that stereotypes work as heuristics (Dovidio, 1999; Gilbert & Hixon, 1991; Macrae & Bodenhausen, 2000). In the presence of a category member, a set of expectations about the properties of this social category appears automatically in people's minds. From this perspective, people in the same context are knowledgeable about the stereotypes in their culture, regardless of their personal opinions or the extent to which they personally hold stereotypes (Beukeboom & Burgers, 2019; Devine, 1989; Lepore & Brown, 1997; Moskowitz et al., 1999; Quadflieg & Macrae, 2011). This theoretical framework, together with the E-Z Reader model, offer a comprehensive account of how stereotype information might affect eye movement control. When encountering stereotype-incongruent information, readers may have difficulty integrating this new information because it violates their world-knowledge. The forward progress of the eyes might thus be interrupted, and attention might shift back to the locus of processing difficulty, i.e., to the stereotype-incongruent information.

6.1.3. Using eye movements to assess stereotype activation

As presented in Chapter 3, stereotypes can operate in unconscious and unintentional ways. Stereotypes are stored in semantic memory in a dormant state until they are activated in a given situation, thereby generating a set of expectancies about the likely traits, behaviors, preferences, or physical features of category members. The activation of stereotypes is considered as automatic because it happens relatively quickly and outside of awareness (Gilbert & Hixon, 1991; Krieglmeyer & Sherman, 2012; Quadflieg & Macrae, 2011). When information violates a stereotype, more cognitive effort is required to integrate this unexpected item into one's mental representation of the text, leading to increased processing difficulty (Banaji & Hardin, 1996a; Bartholow & Dickter, 2008; Kutas & Federmeier, 2000; White et al., 2009).

While there exists an important line of research assessing the automaticity of stereotype activation using response times and event-related brain potentials (e.g., Banaji & Hardin, 1996; Bartholow & Dickter, 2008; Osterhout et al., 1997; Pesciarelli et al., 2019; Proverbio et al., 2018; Rodríguez-Gómez et al., 2020; White et al., 2009), only a few studies used eye-tracking methodologies, and all of them focused on gender stereotypes. In these studies, participants typically read a sentence introducing a role noun (e.g., "The electrician often had good ideas") before reading a second sentence containing an anaphoric pronoun (e.g., "She regularly planned new projects," Reali et al. 2015, see also Duffy & Keir, 2004; Esaulova et al., 2014; Irmen, 2007; Kennison & Trofe, 2003). The critical manipulation is the anaphoric pronoun that can match or not stereotypical expectations about the role noun. Overall, these past studies showed that fixation durations were significantly longer upon and/or after reading a pronoun that led to a mismatch between a role noun and its anaphoric pronoun (e.g., reading "electrician" followed by "she"). These past studies, however, do no offer a comprehensive perspective on how stereotype-incongruent information modulates eye movements. We thus now turn to a broader line of research that assessed the effect of semantic violation during reading.

Research on semantic violation focuses on situations where readers encounter information that violates their expectations and aims to decipher how and when these situations affect readers' eye movements (Warren, 2011). For instance, Ni et al., (1998; see also Braze et al., 2002) presented participants with sentences that either violated world-knowledge as in example (3), or not, as in the baseline condition (4):

(3) It seems that the cats won't usually *bake* the food we put on the porch.

(4) It seems that the cats won't usually *eat* the food we put on the porch.

(Ni et al., 1998, p. 525)

The results revealed that readers made longer first-pass durations when the sentence violated world-knowledge compared to the baseline condition (Braze et al., 2002; Ni et al., 1998). These effects were observed on the region following the critical word (i.e., *bake*), but not on the critical word itself. Regression path times from the critical word, and from the end of the sentence, were also longer in the world-knowledge violation condition compared to the baseline condition (Braze et al., 2002; Ni et al., 1998; for similar findings see Rayner et al., 2004; Warren & McConnell, 2007). Similar studies were conducted to further assess the role of context on the processing of world knowledge violation (Filik, 2008; Warren et al., 2008). While Filik (2008) observed longer first fixation and first-pass durations on the text region following the critical word, Warren et al. (2008) found the effect of world-knowledge violation already on the critical word. Together these past studies offer a detailed and comprehensive description of how and when incongruent information affect eye movement control. Information violating readers' knowledge mostly affected online processing of the critical word or the region following it, leading to increased first fixation, first-pass, and regression durations on and from these regions.

6.1.4. The present study

To assess the specific processing time course of stereotype-incongruent information and presupposed contents, the present study tested two theoretical perspectives from social psychology and pragmatics. On the one hand, cognitive approaches to stereotyping postulate that when information violates a stereotype, more cognitive effort is required to integrate this unexpected item into one's mental representation of a text, resulting in increased processing difficulty (Banaji & Hardin, 1996a; Bartholow & Dickter, 2008; Macrae & Bodenhausen, 2000; J. W. Sherman, Lee, et al., 1998; White et al., 2009). On the other hand, theories in pragmatics posit that presupposed contents (introduced with a definite article) are costly to process when the identification of a salient referent is inconsistent with the encoding of contextual information, i.e., stereotype-incongruent information in the present study (Heim, 1982; C. Roberts, 2003; Schwarz, 2009; Singh et al., 2016; Sperber & Wilson, 1986). Two experiments were thus designed to assess whether readers experience processing difficulties when they encounter stereotype-incongruent information, and whether these difficulties are increased

when such information is presupposed (by the use of definite articles) rather than asserted (by the use of indefinite articles).

The task employed in the next two experiments follows classical research on stereotype reading using eye-tracking methodology (e.g., Duffy & Keir, 2004; Esaulova et al., 2014; Irmen, 2007; Kennison & Trofe, 2003; Reali et al., 2015). Participants read sentences whose content either matched or violated gender (Experiment 1) or nationality-related (Experiment 2) stereotypes. The information was introduced with a definite article, triggering a presupposition of a salient referent, or with an indefinite article, simply introducing a new referent. As previously presented, most past studies assessed the effects of gender stereotypes on reading. To date, only two studies focused the processing of other social categories during reading (with self-paced reading tasks): Dickinson (2011), who analyzed stereotypical inferences regarding heterosexuality, and Lassonde (2015), who assessed stereotypical expectations regarding the behaviors of social groups (e.g., nuns, rockstars). There is thus a need for research investigating stereotypes about other social categories[29].

Participants' gaze was recorded while reading, thus allowing to detect processing difficulties that emerged from the manipulation of stereotype information and article definiteness. Following the E-Z Reader model (Reichle et al., 1998, 2003, 2009) and theoretical accounts of stereotype processing (Sherman et al., 1998; 2013), readers' eye movements should be interrupted when encountering stereotype-incongruent information, and a regression should be initiated as information violates readers' expectations. Specifically, longer first fixation and first-pass durations should be observed from the text region containing stereotype-incongruent information compared to when information confirms a stereotype. As documented in previous studies (e.g., Braze et al., 2002; Ni et al., 1998; Rayner et al., 2004; Warren & McConnell, 2007), these effects might emerge immediately after reading the critical text region, corresponding to spillover effects (Liversedge et al., 1998). The critical text region and its spillover are also expected to elicit longer regression path times when the information violates stereotypical expectations than when it confirms them. In addition, we explored whether stereotype-incongruent information would also increase offline processing (measured by total gaze duration and total sentence processing) compared to stereotype-congruent information. Finally, theoretical approaches in pragmatics (Heim, 1982; C. Roberts, 2003; Schwarz, 2009;

[29] Especially given the fact that gender is considered as a primary social category: Awareness of conventional gender stereotypes has been shown to emerge early, from 3-4 years old (see for e.g., Leinbach et al., 1997; Quinn et al., 2002; Shutts et al., 2009; Weinraub et al., 1984).

Singh et al., 2016; Sperber & Wilson, 1986) predict that processing difficulties should be increased when stereotype-incongruent information is presupposed rather than asserted. For all above-mentioned eye movement measures, longer durations are expected when stereotype-incongruent information is presupposed – through a definite article – than when it is asserted – through an indefinite article.

6.2. Experiment 1

Experiment 1 aimed to further assess the specific time course of processing gender stereotypes and to replicate previous findings with eye-tracking methodology, i.e., that gender stereotype-incongruent information is costly to process. Specifically, online processing difficulties, reflected by increased first fixation, first-pass, and regression durations, were expected from the critical text region containing stereotype-incongruent information and its spillover. These processing difficulties were predicted to be exacerbated when stereotype-incongruent information was presupposed through definite descriptions rather than asserted though indefinite descriptions.

6.2.1. Methods

6.2.1.1. Participants

Participants included 59 French speaking Swiss adults, with normal or corrected-to-normal vision, recruited from the University of Neuchâtel. Only native French speakers were selected to participate in the experiment. The total sample size was set before data collection and based on the sample size estimation for "counterbalanced designs" developed by Westfall et al. (2014). The sample size estimation was conducted using the online app developed by Westfall and colleagues (http://jakewestfall.org/power/). The estimation was run using the "standard case" values of variance components (VPCs; Westfall et al., 2014, p. 2025), with a power set at .90, a medium effect size of $d = 0.50$, and a number of 22 stimuli. The sample size estimation revealed that 58.8 participants were required. No additional participant was recruited once the pre-set sample size of 59 participants was reached. Data from participants who had an accuracy rate for comprehension questions lower than 65% were excluded ($n = 3$). Data from an additional 2 participants had to be excluded due to a loss of tracking signal during the experiment. The final sample size thus resulted in 54 participants (30 women and 24 men; with an age mean of 23.18 years old, $SD = 4.27$).

6.2.1.2. Materials

The stimuli were the same as those employed in the previous study (cf. Chapter 4, Experiment 1). Each stimulus has thus been pre-tested in terms of stereotypicality and plausibility. Stimuli were composed of two sentences. The first sentence introduced the context, whereas the second sentence introduced a specific agent, grammatically marked by gender, which confirmed or violated a gender stereotype related to an occupation. Specifically, the stimuli were created following a 2 x 2 design, manipulating (a) information about the social category, which either matched gender stereotypical expectations or not, and (b) the article introducing the social category, either with a definite article "*le/la*," "the" (presupposition condition), or with an indefinite article "*un/une*," "a/an" (assertion condition). Each stimulus thus varied across four conditions, as in Mari and Müller (2023, Chapter 4, this dissertation): (1) stereotype-congruent and definite NP, (2) stereotype-congruent and indefinite NP, (3) stereotype-incongruent and definite NP, and (4) stereotype-incongruent and indefinite NP (see Table 11). Twenty-two stimuli (half related to female role nouns and half related to male role nouns) and 24 filler sentences were used for this experiment (the list of stimuli and fillers is available in Appendix 1).

Table 11. Example of a stimulus of Experiment 1 in the four experimental conditions.

Sentence	Condition	Stimulus
Context		Pascal est allé contrôler les avions à Genève. Pascal went to verify the planes in Geneva.
Target	Stereotype-congruent and definite NP	[Le pilote] [l'a regardé] [d'un air suspect.] The $\text{pilot}_{\text{male}}$ watched him with suspicion.
	Stereotype-congruent and indefinite NP	[Un pilote] [l'a regardé] [d'un air suspect.] A $\text{pilot}_{\text{male}}$ watched him with suspicion.
	Stereotype-incongruent and definite NP	[La pilote] [l'a regardé] [d'un air suspect.] The $\text{pilot}_{\text{female}}$ watched him with suspicion.
	Stereotype-incongruent and indefinite NP	[Une pilote] [l'a regardé] [d'un air suspect.] A $\text{pilot}_{\text{female}}$ watched him with suspicion.
		[AOI 1] [AOI 2] [AOI 3]

Note. AOIs for the analyses are marked by squared brackets.

6.2.1.3. Apparatus

Eye movements were recorded with a video-based eye-tracking device, SMI RED 5 (SensoMotoric Instruments, Teltow, Germany). To reduce head movements, participants placed their head on a chinrest at a distance of 60 cm from a 22'' screen and the eye-tracking device.

Eye data were recorded binocularly and non-invasively at a sampling rate of 500 Hz. A 5-point calibration was performed before stimuli presentation.

6.2.1.4. Procedure

The experiment was built and run with the SMI Experiment Center software (SensoMotoric Instruments, Teltow, Germany). The purpose of the study was masked to participants and only revealed at the end of the study. Participants were instructed to read the sentences for comprehension.

Before starting the experiment, participants were asked to indicate their age, gender, and mother tongue. A trial started with a fixation cross presented for 500 ms in the middle of the screen. The context and target sentences then appeared together on the screen, written in black with a 32-point Arial font on a light gray background. Three characters represented a visual angle of approximately 1° from a distance of 60 cm to the screen. The stimuli were presented on two lines, with the first line containing the context sentence and the second line containing the target sentence. Participants were instructed to read the sentences for comprehension and to press the space bar to display the next pair of sentences. Participants read only one condition of each stimulus, and as many stimuli from each of the four conditions, resulting in a within-subjects and within-stimuli design (Brauer & Curtin, 2018).

The 22 stimuli and 24 fillers were presented randomly. Comprehension questions were used to assess whether participants were attentive during the whole task. Comprehension questions immediately followed the fillers (no comprehension question was asked about the stimuli). Participants recorded their answers by pressing on the "E" or "I" keys on the keyboard, according to the location of the yes/no answer on the screen. Six practice trials and one comprehension question were used to familiarize participants with the task.

6.2.2. Results

6.2.2.1. Data analysis

Fixations were detected using the SMI built-in algorithm based on velocity, with a threshold set at 40°/s. Fixations shorter than 40 ms were discarded by the algorithm. The recording of eye movement was binocular, but only the data from the dominant eye were analyzed. Eye movement measures were analyzed on three different areas of interests (hereafter AOIs) of the target sentence. The segmentation into AOIs for the analysis is shown in Table

11. AOI 1, 2, and 3 corresponded to the critical region, the spillover region, and the second spillover region, respectively. The two segments following the critical region are traditionally included in the analysis of reading measures because they inform of potential processing difficulties that emerged or persisted after reading a critical region (Liversedge et al., 1998). The three AOIs were created with the SMI software BeGaze Analysis (SensoMotoric Instruments, Teltow, Germany).

Online processing was assessed with the following eye movement measures: first fixation duration, first-pass duration, and regression path time. To assess offline processing, gaze duration and sentence processing were computed (for a detailed description of each eye movement measure, see Table 10). Eye movement measures were logarithmically transformed to meet the assumptions of mixed effects model analyses. The final dataset was composed of 1188 datapoints. Mixed effects modeling was conducted on RStudio (R Core Team, 2019, version 3.6.0) using the lme4 package (Bates, Mächler, et al., 2015).

6.2.2.2. Model selection

Model specification was driven by the experimental design (Barr et al., 2013; Brauer & Curtin, 2018; Winter & Wieling, 2016). Fixed predictors are composed of the interaction between the stereotype condition (stereotype-congruent or stereotype-incongruent) and the NP condition (definite or indefinite article). The study employed a repeated measures design, thus leading to nonindependence in the data for subjects and stimuli. Both were thus included as by-subjects and by-stimuli random intercepts (Brauer & Curtin 2018, p. 401). According to Barr et al. (2013), each fixed predictor that vary within-unit should include a random slope, as well as interactions when all factors vary within-units. In the present study, the stereotype and the NP conditions varied both within-subjects and within-stimuli. Consequently, reading times were assessed with the following maximal mixed effect model: model <- lmer (log eye movement measure ~ stereotype * NP + (stereotype + NP + stereotype*NP | subjects) + (stereotype + NP + stereotype*NP | stimuli)).

The maximal mixed effect model for all eye movement measures of the three analyzed AOIs converged, using the built-in optimization procedures “bobyqa” or “nlminbwrap” of the lme4 package (Bates, Mächler, et al., 2015). All maximal mixed effect models resulted however in a singular fit. Singular fits indicate models’ overparametrization. To balance Type I error rate and statistical power, overparametrized models should be reduced to parsimonious models (Bates, Kliegl, et al., 2015; Bates, Mächler, et al., 2015; Matuschek et al., 2017). Random effect

Principal Component Analysis was thus conducted using the rePCA function of the lme4 package to reach parsimony (Bates, Mächler, et al., 2015). Goodness of fit was estimated with likelihood ratio tests (LRT) and AIC/BIC criteria (Bates, Kliegl, et al., 2015; Matuschek et al., 2017). Table 12 presents the resulting models for all eye movement measures of the three AOIs. The details of model selection and comparison are available in Appendix 2.

Table 12. Parsimonious models for eye movements measures of the AOIs of Experiment 1.

Eye movement measure	Parsimonious model
Online processing	
First fixation duration	
AOI 1	lmer (log first fixation ~ stereotype * NP + (stereotype || subjects) + (NP || stimuli))[1]
AOI 2	lmer (log first fixation ~ stereotype * NP + (1 | subjects) + (NP || stimuli))[1]
AOI 3	lmer (log first fixation ~ stereotype * NP + (1 | subjects) + (0 + NP | stimuli))[2]
First-pass duration	
AOI 1	lmer (log first-pass ~ stereotype * NP + (stereotype + NP + stereotype*NP || subjects) + (stereotype*NP || stimuli))[1]
AOI 2	lmer (log first-pass ~ stereotype * NP + (stereotype*NP || subjects) + (NP || stimuli))
AOI 3	lmer (log first-pass ~ stereotype * NP + (stereotype || subjects) + (stereotype || stimuli))
Regression path time	
AOI 1	lmer (log regression ~ stereotype * NP + (stereotype + NP || subjects) + (1 | stimuli))
AOI 2	lmer (log regression ~ stereotype * NP + (1 | subjects) + (NP || stimuli))[1]
AOI 3	lmer (log regression ~ stereotype * NP + (stereotype + NP || subjects) + (stereotype + NP + stereotype*NP || stimuli))[1]
Offline processing	
Total gaze duration	
AOI 1	lmer (log total gaze ~ stereotype * NP + (stereotype + NP + stereotype*NP || subjects) + (1 | stimuli))
Sentence processing	lmer (log processing ~ stereotype * NP + (NP || subjects) + (stereotype*NP || stimuli))

Note. Parsimonious models selected after random effect Principal Component Analysis and estimation of goodness of fit with likelihood ratio tests and AIC/BIC criteria.
Optimizer employed: [1] bobyqa, [2] nlminbwrap.

6.2.2.3. Online processing measures

First fixation durations. First fixation durations were expected to be longer on AOIs 1, 2, and 3 in stereotype-incongruent conditions than in stereotype-congruent conditions. Definite descriptions were expected to exacerbate processing difficulties of stereotype-incongruent information, thus leading to longer first fixation duration in this condition than in

the others. The analysis of first fixation durations for AOI 1 and AOI 2, however, revealed no significant effect of the stereotype condition, of the NP condition, or of their interaction (all *p*s > .14, see Table 13). However, a significant effect of the NP condition was observed on AOI 3, $t(272.15) = -2.1$, $p = .036$. First fixation durations were longer when stereotype-incongruent information was introduced with a definite article ($M = 551.44$ ms, $SD = 387.37$) compared to with an indefinite article ($M = 509.79$ ms, $SD = 345.27$). No effect of the stereotype condition or of the interaction was observed (all *p*s > .40; see Table 13).

First-pass durations. Similarly, first-pass durations were expected to be longer on AOIs 1, 2, and 3 in stereotype-incongruent conditions than in stereotype-congruent conditions, with increased processing difficulties when information was introduced with definite descriptions. The analysis of first-pass duration for the three AOIs, however, revealed no significant effect of the stereotype condition, of the NP condition, or of their interaction (all *p*s > .30, see Table 13).

Regression path times. When readers encounter information that violates their expectations, forward movement is cancelled in favor of regressions to the unexpected element. Regression path times were thus expected to be longer on AOIs 1, 2, and 3 in stereotype-incongruent conditions than in stereotype-congruent conditions, with increased processing difficulties when information was introduced with definite descriptions. The analysis of regression path times for AOI 1 and AOI 3 revealed no significant effect of the three predictors (all *p*s > .42; see Table 13). However, a significant effect of the stereotype condition was observed on AOI 2, with longer regression times in the stereotype-incongruent condition ($M = 546.36$ ms, $SD = 643.98$) compared to the stereotype-congruent condition ($M = 459.02$ ms, $SD = 508.3$), $t(1048.7) = -2.04$, $p = .042$. No effect of the NP condition or of the interaction was observed (all *p*s > .52; see Table 13).

Table 13. Statistical results of the parsimonious models for online processing measures of AOI 1, 2 and 3 in Experiment 1.

Fixed effects							**Random effects**		
	Estimate	SE	CI (95%)	*t*-value	DF	*p*-value		Var.	SD
First fixation duration (AOI 1)									
(Intercept)	5.21	0.04	[5.13, 5.28]	136.69	125.54	<.001	Subjects intercept	0.029	0.17
Stereotype - Congruent	0.01	0.05	[-0.08, 0.10]	0.30	214.79	.763	Subjects slope (stereotype)	0.019	0.14
NP - Indefinite	-0.06	0.04	[-0.14, 0.02]	-1.46	200.24	.147	Stimuli intercept	0.0002	0.02
Interaction	0.02	0.06	[-0.09, 0.14]	0.35	1037.01	.730	Stimuli slope (NP)	0.001	0.02
First fixation duration (AOI 2)									
(Intercept)	5.15	0.04	[5.08, 5.23]	130.37	120.50	<.001	Subjects intercept	0.034	0.18
Stereotype - Congruent	-0.008	0.04	[-0.08, 0.07]	-0.19	1055.03	.850	Stimuli intercept	0.004	0.07
NP - Indefinite	0.002	0.04	[-0.07, 0.08]	0.04	153.46	.966	Stimuli slope (NP)	0.0004	0.02
Interaction	-0.04	0.05	[-0.15, 0.06]	-0.77	1057.28	.443			
First fixation duration (AOI 3)									
(Intercept)	5.22	0.03	[5.15, 5.28]	149.65	225.39	<.001	Subjects intercept	0.02	0.14
Stereotype - Congruent	-0.02	0.04	[-0.10, 0.06]	-0.60	1089.49	.548	Stimuli slope (NP)	0.001	0.03
NP - Indefinite	-0.09	0.04	[-0.17, -0.005]	-2.10	272.15	**.036**			
Interaction	0.04	0.06	[-0.07, 0.16]	0.77	1091.68	.439			
First-pass duration (AOI 1)									
(Intercept)	5.54	0.05	[5.44, 5.64]	11.31	92.05	<.001	Subjects intercept	0.05	0.22
Stereotype - Congruent	0.01	0.05	[-0.08, 0.10]	0.30	150.28	.765	Subjects slope (stereotype)	0.02	0.14
NP - Indefinite	0.05	0.04	[-0.04, 0.13]	1.03	192.69	.303	Subjects slope (NP)	0.002	0.04
Interaction	-0.03	0.07	[-0.16, 0.09]	-0.47	101.98	.638	Subjects slope (interaction)	0.01	0.10
							Stimuli intercept	0.01	0.11
							Stimuli slope (interaction)	0.005	0.07

Table 13 (*continued*).

Fixed effects							**Random effects**		
	Estimate	SE	CI (95%)	*t*-value	DF	*p*-value		Var.	SD
First-pass duration (AOI 2)									
(Intercept)	5.38	0.04	[5.29, 5.47]	123.92	115.59	<.001	Subjects intercept	0.04	0.21
Stereotype - Congruent	-0.001	0.04	[-0.08, 0.08]	-0.03	1002.13	.975	Subjects slope (interaction)	0.005	0.07
NP - Indefinite	0.02	0.04	[-0.06, 0.11]	0.52	130.07	.603	Stimuli intercept	0.004	0.07
Interaction	-0.03	0.06	[-0.15, 0.08]	-0.59	388.26	.556	Stimuli slope (NP)	0.003	0.06
First-pass duration (AOI 3)									
(Intercept)	5.87	0.07	[5.72, 6.03]	76.89	53.91	<.001	Subjects intercept	0.06	0.25
Stereotype - Congruent	-0.03	0.06	[-0.15, 0.09]	-0.40	65.65	.689	Subjects slope (stereotype)	0.01	0.10
NP - Indefinite	0.02	0.05	[-0.09, 0.13]	0.31	1026.75	.758	Stimuli intercept	0.06	0.26
Interaction	0.02	0.08	[-0.13, 0.17]	0.27	1025.61	.782	Stimuli slope (stereotype)	0.01	0.11
Regression path time (AOI 1)									
(Intercept)	5.88	0.05	[5.76, 5.99]	103.05	98.06	<.001	Subjects intercept	0.08	0.28
Stereotype - Congruent	-0.02	0.06	[-0.15, 0.09]	-0.37	161.65	.711	Subjects slope (stereotype)	0.05	0.23
NP - Indefinite	-0.03	0.05	[-0.14, 0.08]	-0.47	198.71	.639	Subjects slope (NP)	0.02	0.13
Interaction	0.02	0.07	[-0.12, 0.17]	0.27	1009.89	.784	Stimuli intercept	0.008	0.09
Regression path time (AOI 2)									
(Intercept)	5.92	0.05	[5.81, 6.03]	102.55	126.37	<.001	Subjects intercept	0.04	0.21
Stereotype - Congruent	-0.13	0.06	[-0.25, -0.004]	-2.04	1048.67	**.042**	Stimuli intercept	0.01	0.10
NP - Indefinite	-0.03	0.07	[-0.16, 0.09]	-0.57	143.09	.570	Stimuli slope (NP)	0.005	0.07
Interaction	0.06	0.09	[-0.12, 0.23]	0.64	1051.06	.522			
Regression path time (AOI 3)									
(Intercept)	6.26	0.07	[6.11, 6.41]	82.38	63.89	<.001	Subjects intercept	0.09	0.29
Stereotype - Congruent	-0.01	0.07	[-0.15, 0.12]	-0.18	57.60	.855	Subjects slope (stereotype)	0.01	0.11
NP - Indefinite	0.002	0.06	[-0.13, 0.14]	0.03	52.81	.978	Subjects slope (NP)	0.009	0.09
Interaction	0.07	0.09	[-0.11, 0.26]	0.79	72.98	.429	Stimuli intercept	0.04	0.21
							Stimuli slope (stereotype)	0.02	0.13
							Stimuli slope (NP)	0.008	0.09
							Stimuli slope (interaction)	0.02	0.14

Note. DF stands for degrees of freedom, SE for standard error, CI for confidence interval, Var. for Variance, and SD for standard deviation. Values in bold are significant at $p < .05$ (calculated with Satterthwaites approximations).

6.2.2.4. Offline processing measures

Total gaze durations. The text region containing stereotype information (i.e., AOI 1) should receive more attention, as reflected by longer total gaze durations in stereotype-incongruent compared to stereotype-congruent conditions. However, the analysis of total gaze durations on AOI 1 revealed no significant effect of the three predictors (all *p*s > .40, see Table 14).

Sentence processing. Overall, sentence processing was expected to be longer in stereotype-incongruent than in stereotype-congruent conditions. However, no effect of the three predictors was observed for processing times of the whole sentence (all *p*s > .20, see Table 14).

Table 14. Statistical results of the parsimonious models for offline processing measures in Experiment 1.

Offline processing measures						
Total gaze duration (AOI 1)						
	Estimate	SE	CI (95%)	*t*-value	DF	*p*-value
Fixed effects						
(Intercept)	6.15	0.07	[6.22, 6.28]	91.99	85.89	<.001
Stereotype - Congruent	-0.04	0.05	[-0.14, 0.06]	-0.84	140.66	.405
NP - Indefinite	0.02	0.05	[-0.07, 0.12]	0.44	179.35	.657
Interaction	-0.03	0.06	[-0.16, 0.10]	-0.42	246.79	.674
Random effects	Var.	SD				
Subjects intercept	0.12	0.35				
Subjects slope (stereotype)	0.03	0.17				
Subjects slope (NP)	0.009	0.09				
Subjects slope (interaction)	0.002	0.05				
Stimuli intercept	0.02	0.15				
Sentence processing						
	Estimate	SE	CI (95%)	*t*-value	DF	*p*-value
Fixed effects						
(Intercept)	8.66	0.05	[8.57, 8.76]	178.41	84.04	<.001
Stereotype - Congruent	-0.04	0.02	[-0.08, 0.02]	-1.22	1051.83	.220
NP - Indefinite	0.01	0.03	[-0.04, 0.06]	0.53	173.49	.594
Interaction	-0.0006	0.04	[-0.07, 0.07]	-0.02	157.08	.986
Random effects	Var.	SD				
Subjects intercept	0.08	0.28				
Subjects slope (NP)	0.005	0.07				
Stimuli intercept	0.01	0.11				
Stimuli slope (interaction)	0.001	0.03				

Note. DF stands for degrees of freedom, SE for standard error, CI for confidence interval, Var. for Variance, and SD for standard deviation. Values in bold are significant at $p < .05$ (calculated with Satterthwaites approximations).

6.2.3. Discussion

Experiment 1 provided small support in favor of increased processing costs when reading stereotype-incongruent information. Specifically, reading was significantly longer when it contradicted a gender stereotype for only one measure of online processing, namely regression path times. Regression path time is an interesting measure as it considers any return to previously read regions before resuming reading. That is, after reading the critical region (e.g., *le/la pilote*; the $\text{pilot}_{\text{male/female}}$), participants took more time to re-read the previous regions when they encountered a stereotype-incongruent compared to stereotype-congruent information. This finding is consistent with theoretical accounts of eye movement control during reading (Reichle et al., 1998, 2003, 2009) and cognitive approaches to stereotype processing (Sherman et al., 1998; 2013): When encountering gender stereotype-incongruent information, participants interrupted their forward reading to return to the locus of processing difficulty.

Experiment 1 also investigated whether linguistic markers could increase the costs of processing gender stereotype-incongruent information. The findings of Experiment 1 showed that first fixation durations were longer on the second spillover when stereotype-incongruent information was introduced with a definite article (e.g., *la pilote,* the $\text{pilot}_{\text{female}}$) compared to an indefinite article (e.g., *une pilote*, a $\text{pilot}_{\text{female}}$). This is in line with theoretical approaches in pragmatics (Heim, 1982; Roberts, 2003; Schwarz, 2009; Singh et al., 2016; Sperber & Wilson, 1986) and replicates previous findings showing that processing definite compared to indefinite NPs is significantly costlier when information violates a gender stereotype (Mari & Müller, 2023, Chapter 4, this dissertation). The processing of definite articles was costly, because information that contradicted gender stereotypes made it difficult to identify the referent from the preceding context sentence.

Regarding the exploratory analyses of offline processing measures, no effect of the stereotype information or article was observed. The analyses of the text region and the whole sentence containing gender stereotype information revealed that readers did not show differential attention allocation based on the experimental condition. Rather, information contradicting gender stereotypes only affected the first reading of the text, but not its later reading. This finding suggests that participants' gender stereotypical expectations about occupations activated while reading, and influenced their online processing of the written text.

However, stereotype activation did not affect later processing. This point is further addressed in the General Discussion.

6.3. Experiment 2

Experiment 2 aimed to further assess the specific time course of processing nationality stereotypes and to replicate previous findings (i.e., that nationality stereotype-incongruent information is costly to process) with eye-tracking methodology. As for Experiment 1, online processing difficulties, reflected by increased first fixation, first-pass, and regression durations, were expected from the critical text region containing stereotype-incongruent information and its spillover. These processing difficulties were predicted to be exacerbated when stereotype-incongruent information was presupposed rather than asserted. The structure of the stimuli in Experiment 2 also allowed to exclusively assess whether definite articles were more quickly read than indefinite articles in redundant contexts.

6.3.1. Methods

6.3.1.1. Participants

Participants included 49 native French speaking Swiss adults, with normal or corrected-to-normal vision, recruited from the University of Neuchâtel. The total sample size was estimated for "counterbalanced designs" using the online app (http://jakewestfall.org/power/) developed by Westfall et al. (2014) with "standard case" values of variance components, a power of .85, a medium effect size of $d = 0.50$, and a number of 20 stimuli. The sample size estimation revealed that 48.8 participants were required. No additional participant was recruited once the pre-set sample size of 49 participants was reached. As for Experiment 1, we excluded data from participants who had an accuracy rate for comprehension questions lower than 65% ($n = 4$). The final sample size resulted 45 participants (26 women and 19 men; with an age mean of 23.49 years old, $SD = 4.25$).

6.3.1.2. Materials

The stimuli were the same as those employed in the previous study (cf. Chapter 4, Experiment 2) and were thus pre-tested in terms of stereotypicality. Each stimulus was composed of two sentences: The first sentence introduced the context, whereas the second sentence introduced representatives of specific countries, displaying a behavior or properties

that matched or violated a nationality stereotype. Specifically, the stimuli were created following a 2 x 2 design, manipulating (a) information about the social category, which either confirmed or violated nationality stereotypical expectations, and (b) the article introducing the social category, either definite (working as a presupposition, favoring a generalized and taxonomic representation of the social category) or indefinite (working as an assertion, favoring a single occurrence reading of the stereotype). Each stimulus thus varied across four conditions, as in Mari and Müller (2023, Chapter 4, this dissertation): (1) stereotype-congruent and definite NP, (2) stereotype-congruent and indefinite NP, (3) stereotype-incongruent and definite NP, and (4) stereotype-incongruent and indefinite NP (see Table 15). In addition to the 20 stimuli, 24 filler sentences were used to hide the purpose of the study (the list of stimuli and fillers is available in Appendix 3).

Table 15. Example of a stimulus of Experiment 2 in the four experimental conditions.

Sentence	Condition	Stimulus
Context		Nathan est allé en Angleterre / France l'hiver dernier. Nathan went to England / France last winter.
Target	Stereotype-congruent and definite NP	[Les Anglais] [lui ont semblé] [particulièrement polis] [durant son voyage.] The English seemed particularly polite during his stay.
	Stereotype-congruent and indefinite NP	[Un Anglais] [lui a semblé] [particulièrement poli] [durant son voyage.] An English seemed particularly polite during his stay.
	Stereotype-incongruent and definite NP	[Les Français] [lui ont semblé] [particulièrement polis] [durant son voyage.] The French seemed particularly polite during his stay.
	Stereotype-incongruent and indefinite NP	[Un Français] [lui a semblé] [particulièrement poli] [durant son voyage.] A French seemed particularly polite during his stay.
		[AOI 1] [AOI 2] [AOI 3]

Note: AOIs for the analyses are marked by squared brackets.

6.3.1.3. Apparatus

The apparatus was the same as the one described in Experiment 1.

6.3.1.4. Procedure

The procedure was the same as the one described in Experiment 1.

6.3.2. Results

6.3.2.1. Data analysis

Fixations were detected using the SMI built-in algorithm as in Experiment 1 (with a velocity threshold set at 40°/s and minimum fixation duration of 40 ms). The recording of eye movement was binocular, but only the data from the dominant eye were analyzed.

Four AOIs were considered for the analysis: (1) the text region containing the definite/indefinite NP that introduced the social category (i.e., inhabitants of a country), (2) its spillover region to assess potential persistence of processing difficulty, (3) the region presenting stereotype-congruent/incongruent information, and (4) its spillover region. The segmentation into AOIs for the analysis is shown in Table 15. Again, six eye movement measures were computed to assess online and offline processing (cf. Table 10). Eye movement measures were logarithmically transformed to meet the assumptions of mixed effects model analyses. The final dataset was composed of 980 datapoints. Mixed effects modeling was conducted on RStudio (R Core Team, 2019, version 3.6.0) using the lme4 package (Bates et al., 2015a).

6.3.2.2. Model selection

Model specification was based on the same procedure as described in Experiment 1 (Section 6.2.3.2) As illustrated in Table 15, information confirming/contradicting nationality stereotypes is only introduced at AOI 3. Online processing measures for AOIs 1 and 2 could thus only be affected by the NP condition, but not by the stereotype condition, whereas AOIs 3 and 4 could be affected by both conditions. For this reason, two models were used to analyze eye-movement measures for online processing on the different AOIs:

1) AOI 1 and AOI 2: model1 <- lmer (log online measures ~ NP + (NP | subjects) + (NP | stimuli))
2) AOI 3 and AOI 4: model2 <- lmer (log online measures ~ stereotype * NP + (stereotype + NP + stereotype*NP | subjects) + (stereotype + NP + stereotype*NP | stimuli))

Eye movement measures for offline processing, however, could be affected by the two conditions, and were thus assessed with the following maximal mixed effect model: model3 <- lmer (log offline measures ~ stereotype * NP + (stereotype + NP + stereotype*NP | subjects) + (stereotype + NP + stereotype*NP | stimuli)).

The maximal mixed effect model for all eye movement measures of the three analyzed AOIs converged, using the built-in optimization procedures "bobyqa" of the lme4 package (Bates, Mächler, et al., 2015). Some maximal mixed effect models, however, resulted in a singular fit. Singular fits indicate that the model is overparametrized. To balance Type I error rate and statistical power, overparametrized models should be reduced to parsimonious models (Bates, Kliegl, et al., 2015; Bates, Mächler, et al., 2015; Matuschek et al., 2017). Random effect Principal Component Analysis was thus conducted using the rePCA function of the lme4 package to reach parsimony (Bates, Mächler, et al., 2015). Goodness of fit was estimated with likelihood ratio tests (LRT) and AIC/BIC criteria (Bates, Kliegl, et al., 2015; Matuschek et al., 2017). Table 16 presents the resulting models for all eye movement measures of the three AOIs. The details of model selection and comparison are available in Appendix 4.

Table 16. Parsimonious models for eye movements measures of the AOIs of Experiment 2.

Eye movement measure	**Parsimonious model**
Online processing	
First fixation duration	
AOI 1	lmer (log first fixation ~ NP + (NP || subjects) + (1 | stimuli))[1]
AOI 2	lmer (log first fixation ~ NP + (NP || subjects))[1]
AOI 3	lmer (log first fixation ~ stereotype * NP + (NP + stereotype * NP || subjects) + (1 | stimuli))[1]
AOI 4	lmer (log first fixation ~ stereotype * NP + (stereotype + stereotype * NP || subjects) + (1 | stimuli))[1]
First-pass duration	
AOI 1	lmer (log first-pass ~ NP + (NP || subjects) + (1 | stimuli))[1]
AOI 2	lmer (log first-pass ~ NP + (NP | subjects) + (NP | stimuli))
AOI 3	lmer (log first-pass ~ stereotype * NP + (NP || subjects) + (1 | stimuli))
AOI 4	lmer (log first-pass ~ stereotype * NP + (stereotype || subjects) + (stereotype * NP || stimuli))
Regression path time	
AOI 3	lmer (log regression ~ stereotype * NP + (NP || subjects) + (1 | stimuli))
AOI 4	lmer (log regression ~ stereotype * NP + (stereotype || subjects) + (NP || stimuli))
Offline processing	
Total gaze duration	
AOI 1	lmer (log total gaze ~ stereotype * NP + (NP || subjects) + (NP || stimuli))
AOI 3	lmer (log total gaze ~ stereotype * NP + (NP || subjects) + (NP || stimuli))
Sentence processing	lmer (log processing time ~ stereotype * NP + (NP || subjects) + (stereotype + NP || stimuli))

Note: Parsimonious models selected after random effect Principal Component Analysis and estimation of goodness of fit with likelihood ratio tests and AIC/BIC criteria.
[1] Optimizer bobyqa employed.

6.3.2.3. Online processing measures

First fixation durations. First fixation durations on AOIs 3 and 4 were expected to be longer in stereotype-incongruent conditions than in stereotype-congruent conditions. Moreover, following theories in pragmatics, in redundant contexts, first fixation durations on AOIs 1 and 2 should be shorter when the content is presupposed through definite descriptions than asserted through indefinite descriptions. The analysis of first fixation durations for the four AOIs however revealed no significant effect of the stereotype condition, of the NP condition, or of their interaction (all *p*s > .10; cf. Table 17).

First-pass durations. Similarly, first-pass durations on AOI 3 and 4 were expected to be longer in stereotype-incongruent than in stereotype-congruent conditions. First-pass durations on AOI 1 and 2 were expected to be shorted for definite than indefinite descriptions due to context redundance. However, the analysis of first-pass durations on all AOIs revealed no significant effect of the three predictors (all *p*s > .12; cf. Table 17).

Regression path times. When information violates readers' expectations, forward saccades are cancelled in favor of regressions to the unexpected element. Regression path times were thus expected to be longer on AOIs 3 and 4 in stereotype-incongruent than in stereotype-congruent conditions, with increased processing difficulties when information was introduces with definite descriptions. For AOI 3, no effect of the three predictors was observed (all *p*s > .16; cf. Table 17). A marginal effect of the stereotype condition on AOI 4 was nonetheless observed: regression path times were longer when information contradicted a stereotype (M = 775.11 ms, SD = 665.38) compared to when it confirmed a stereotype (M = 724.10 ms, SD = 588.61), $t(195.19) = -1.89$, $p = .060$. No other effect was observed for AOI 4 (cf. Table 17).

6.3.2.4. Offline processing measures

Total gaze durations. Text regions containing social category and stereotype information, i.e., AOI 1 and AOI 3, were expected to receive longer total gaze durations when information was incongruent with stereotypical expectations than when it was congruent. However, the analysis of total gaze durations for AOIs 1 and 3 revealed no significant effect of the three predictors (all *p*s > .09; cf. Table 18).

Sentence processing. Overall, sentence processing was expected to be longer in stereotype-incongruent than in stereotype-congruent conditions. The analysis of sentence processing revealed no significant effect of the three predictors (all *p*s > .30; cf. Table 18).

Table 17. Statistical results of the parsimonious models for online processing measures of each AOI in Experiment 2.

Fixed effects							**Random effects**		
	Estimate	SE	CI (95%)	*t*-value	DF	*p*-value		Var.	SD
First fixation duration (AOI 1)									
(Intercept)	5.18	0.03	[5.11, 5.25]	162.03	51.47	<.001	Subjects intercept	0.016	0.13
NP - Indefinite	0.007	0.03	[-0.06, 0.07]	0.22	64.79	.830	Subjects slope (NP)	0.001	0.03
							Stimuli intercept	0.002	0.05
First fixation duration (AOI 2)									
(Intercept)	5.11	0.03	[5.05, 5.16]	172.01	63.02	<.001	Subjects intercept	0.018	0.13
NP - Indefinite	0.02	0.04	[-0.05, 0.09]	0.66	64.29	.513	Subjects slope (NP)	0.011	0.11
First fixation duration (AOI 3)									
(Intercept)	5.04	0.04	[4.95, 5.12]	112.31	114.43	<.001	Subjects intercept	0.036	0.19
Stereotype - Congruent	0.07	0.05	[-0.01, 0.16]	1.65	752.44	.100	Subjects slope (NP)	0.003	0.06
NP - Indefinite	0.03	0.05	[-0.06, 0.12]	0.62	188.72	.535	Subjects slope (interaction)	0.006	0.08
Interaction	-0.04	0.06	[-0.17, 0.09]	-0.58	285.33	.560	Stimuli intercept	0.002	0.05
First fixation duration (AOI 4)									
(Intercept)	5.20	0.05	[5.10, 5.30]	104.54	99.46	<.001	Subjects intercept	0.05	0.22
Stereotype - Congruent	0.03	0.05	[-0.07, 0.13]	0.55	143.79	.583	Subjects slope (stereotype)	0.02	0.14
NP - Indefinite	-0.04	0.05	[-0.13, 0.06]	-0.77	754.11	.443	Subjects slope (interaction)	0.004	0.06
Interaction	0.07	0.07	[-0.07, 0.20]	0.96	248.01	.336	Stimuli intercept	0.003	0.05
First-pass duration (AOI 1)									
(Intercept)	5.54	0.04	[5.47, 5.62]	151.02	55.31	<.001	Subjects intercept	0.04	0.19
NP - Indefinite	0.001	0.03	[-0.06, 0.06]	0.04	57.87	.968	Subjects slope (NP)	0.006	0.08
							Stimuli intercept	0.002	0.04

Table 17 (*continued*).

Fixed effects							**Random effects**			
	Estimate	SE	CI (95%)	*t*-value	DF	*p*-value		Var.	SD	Corr.
First-pass duration (AOI 2)										
(Intercept)	5.44	0.05	[5.34, 5.54]	112.26	35.40	<.001	Subjects intercept	0.04	0.21	
NP - Indefinite	-0.06	0.04	[-0.16, 0.03]	-1.45	16.96	.166	Subjects slope (NP)	0.004	0.06	-0.06
							Stimuli intercept	0.01	0.12	
							Stimuli slope (NP)	0.01	0.11	-0.71
First-pass duration (AOI 3)										
(Intercept)	5.73	0.06	[5.60, 5.86]	87.65	115.69	<.001	Subjects intercept	0.06	0.26	
Stereotype - Congruent	0.11	0.07	[-0.03, 0.24]	1.55	795.38	.123	Subjects slope (NP)	0.01	0.11	
NP - Indefinite	-0.03	0.07	[-0.17, 0.10]	-0.47	220.44	.638	Stimuli intercept	0.008	0.09	
Interaction	-0.04	0.09	[-0.22, 0.15]	-0.40	790.95	.688				
First-pass duration (AOI 4)										
(Intercept)	5.77	0.06	[5.65, 5.89]	97.13	96.18	<.001	Subjects intercept	0.07	0.27	
Stereotype - Congruent	0.02	0.06	[-0.11, 0.14]	0.26	183.57	.795	Subjects slope (stereotype)	0.02	0.12	
NP - Indefinite	-0.04	0.06	[-0.15, 0.08]	-0.62	773.50	.539	Stimuli intercept	0.002	0.04	
Interaction	0.10	0.09	[-0.07, 0.27]	1.17	154.66	.254	Stimuli slope (interaction)	0.01	0.11	
Regression path time (AOI 3)										
(Intercept)	6.56	0.08	[6.40, 6.73]	79.73	82.02	<.001	Subjects intercept	0.09	0.30	
Stereotype - Congruent	-0.07	0.07	[-0.21, 0.07]	-0.95	791.24	.343	Subjects slope (NP)	0.03	0.16	
NP - Indefinite	-0.11	0.08	[-0.26, 0.04]	-1.40	184.99	.162	Stimuli intercept	0.04	0.20	
Interaction	0.13	0.10	[-0.07, 0.33]	1.29	788.06	.197				
Regression path time (AOI 4)										
(Intercept)	6.38	0.06	[6.25, 6.50]	98.53	89.73	<.001	Subjects intercept	0.04	0.20	
Stereotype - Congruent	-0.14	0.07	[-0.28, 0.005]	-1.89	195.19	*.060*	Subjects slope (stereotype)	0.03	0.17	
NP - Indefinite	-0.05	0.07	[-0.19, 0.09]	-0.72	116.57	.471	Stimuli intercept	0.02	0.13	
Interaction	0.18	0.09	[-0.01, 0.37]	1.60	770.24	.109	Stimuli slope (NP)	0.006	0.07	

Note. DF stands for degrees of freedom, SE for standard error, CI for confidence interval, Var. for Variance, SD for standard deviation and Corr. for correlation. Values in bold are significant at $p < .05$ (calculated with Satterthwaites approximations). Values in italics are marginally significant.

Table 18. Statistical results of the parsimonious models for offline processing measures in Experiment 2.

Fixed effects							**Random effects**		
	Estimate	SE	CI (95%)	*t*-value	DF	*p*-value		Var.	SD
Total fixation duration (AOI 1)									
(Intercept)	5.87	0.06	[5.76, 5.98]	104.12	86.53	<.001	Subjects intercept	0.07	0.26
Stereotype - Congruent	0.04	0.05	[-0.07, 0.15]	0.75	748.22	.453	Subjects slope (NP)	0.003	0.06
NP - Indefinite	0.09	0.06	[-0.02, 0.21]	1.69	84.99	.096	Stimuli intercept	0.002	0.05
Interaction	-0.01	0.08	[-0.16, 0.14]	-0.16	751.59	.871	Stimuli slope (NP)	0.002	0.05
Total fixation duration (AOI 3)									
(Intercept)	6.63	0.08	[6.47, 6.78]	86.43	69.41	<.001	Subjects intercept	0.11	0.33
Stereotype - Congruent	-0.005	0.05	[-0.10, 0.09]	-0.10	775.10	.918	Subjects slope (NP)	0.002	0.05
NP - Indefinite	0.05	0.05	[-0.05, 0.16]	0.96	61.78	.342	Stimuli intercept	0.04	0.21
Interaction	-0.05	0.07	[-0.18, 0.09]	-0.74	769.57	.462	Stimuli slope (NP)	0.005	0.07
Sentence processing									
(Intercept)	8.74	0.05	[8.65, 8.83]	191.72	63.49	<.001	Subjects intercept	0.07	0.26
Stereotype - Congruent	-0.01	0.03	[-0.07, 0.04]	-0.54	82.38	.591	Subjects slope (NP)	0.0007	0.03
NP - Indefinite	0.03	0.03	[-0.03, 0.08]	1.03	69.91	.305	Stimuli intercept	0.003	0.06
Interaction	-0.003	0.04	[-0.08, 0.07]	-0.07	789.51	.941	Stimuli slope (stereotype)	0.0005	0.02
							Stimuli slope (NP)	0.0004	0.02

Note. DF stands for degrees of freedom, SE for standard error, CI for confidence interval, Var. for Variance, and SD for standard deviation. Values in bold are significant at $p < .05$ (calculated with Satterthwaites approximations).

6.3.3. Discussion

Experiment 2 failed to provide evidence for increased processing costs when reading information that contradicts nationality stereotypes. No difference was observed for first fixation and first-pass durations between conditions. Only a marginal effect of nationality stereotypes was observed for regression path times. That is, upon reaching the spillover region, participants made longer regressions to previous regions when information contradicted stereotypical expectations than when it confirmed them. The exploratory analyses of offline processing measures did not show any effect of stereotype information. Readers did not allocate more attention to text regions, or sentences, that contained stereotype-incongruent information compared to stereotype-congruent information. While the marginal effect observed for regression durations is consistent with Experiment 1's findings, the results of Experiment 2 generally failed to replicate Mari and Müller's study (2023, Chapter 4, this dissertation). This issue is further discussed in the General Discussion.

Experiment 2 also allowed to exclusively assess whether definite articles are more quickly read than indefinite articles in redundant contexts. However, no difference was observed between the conditions. The current experiment provides no evidence that participant's reading was facilitated by definite descriptions. This is contradictory with the results of Mari and Müller (2023, Chapter 4, this dissertation) and theoretical accounts in pragmatics (Heim, 1982; Roberts, 2003; Schwarz, 2009; Singh et al., 2016; Sperber & Wilson, 1986), positing that presupposed contents are faster to process than asserted ones in redundant contexts. These null results might be due to the methodology employed in the present study. Most past research assessing the processing of presupposed contents employed self-paced reading tasks (e.g., Clifton, 2013; Garrod & Sanford, 1977; Haviland & Clark, 1974; Mari & Müller, 2023; Müller & Mari, 2021; Singh et al., 2016). In the present study, participants reading pace was measured in almost natural settings, while reading on a computer. In such settings, reading pace is half more rapid than in self-paced reading tasks and word skipping is frequent (Rayner, 1998). Importantly, previous studies showed that, in normal reading settings, function words are less fixated than content words, and words of 2-3 letters are skipped 75% of the time (Carpenter & Just, 1983; Rayner & Duffy, 1986; Rayner & McConkie, 1976; Vitu et al., 1995). For this reason, it is most likely that in/definite articles in the present study were skipped, thereby leading to inconclusive results. Consequently, eye-tracking methodology revealed to be inadequate for assessing subtle effects of linguistic markers of definiteness on information processing.

6.4. General discussion

The current research investigated two theoretical perspectives in social cognition and pragmatics proposing that stereotypical expectations and specific linguistic markers modulate information processing. Using eye-tracking methodology, the present study assessed the specific processing time course of stereotype-incongruent information and presupposed contents. More concretely, two experiments assessed whether readers experienced processing difficulties when they encountered information that contradicted gender and nationality stereotypes, and whether these difficulties were increased when such information was presupposed (through definite descriptions) rather than asserted (through indefinite descriptions).

The results of Experiment 1 revealed that participants slowed down and interrupted their forward reading of the text after encountering *gender* stereotype-incongruent information. Experiment 2, however, failed to demonstrate significant effects of *nationality* stereotype-incongruent information: Only a marginal effect of stereotype information was observed on regression durations. This finding contrasts somewhat with the results of Mari and Müller (2023, Chapter 4 this dissertation), that showed significant effects of both gender and nationality stereotypes on information processing. However, compared to Mari and Müller, the present study assessed participants' reading in natural settings, rather than with a self-paced reading task. Importantly, self-paced reading tasks increase cognitive load due to the task's demands (Rayner, 1998) and might have enhanced the processing costs of stereotype-incongruent information. Indeed, previous research clearly established the influence of task demands and cognitive load on the activation of stereotypical expectations (Dovidio, 1999; Fiske & Taylor, 2013; Gilbert & Hixon, 1991). As such, processing difficulties might have been exacerbated because of the cognitive load imposed on participants with self-paced reading tasks. In contrast, the present study adds a novel perspective to the processing of stereotype-incongruent information, showing that in natural settings, readers made longer regressions for items that violated their expectations about social categories, especially about gender, a primary social category.

Interestingly, in both the self-paced reading study (Chapter 4) and the eye-tracking study (this chapter), reading times were affected at the segment following the violation, rather than at the critical segment. One possible explanation for the delayed effect of the violation pertains to the grammatical analyses of noun phrases. Past studies using eye-tracking measures to assess

the processing of stereotype information found that reading times were significantly slowed down *after* reading the critical region, when a noun phrase was used (e.g., after reading "this woman" when it is preceded by "electrician," Esaulova et al., 2014, see also Irmen, 2007). In contrast, studies using pronominal anaphors (e.g., "she" preceded by "electrician," Reali et al., 2015) observed effects of the violation *immediately upon* reading the pronoun (Esaulova et al., 2014; Kennison & Trofe, 2003). These effects can be explained by the fact that noun phrases, unlike pronouns, do not require an immediate interpretation of the antecedent to be comprehended (Garrod & Sanford, 1995). The interpretation of the critical noun phrases was thus probably delayed by the additional semantic content that needed to be processed (cf. Esaulova et al., 2014).

A second explanation pertains to the fact that processing stereotype-incongruent information is similar to processing semantic violation. Studies assessing semantic violation with eye-tracking measures also observed increased reading times *after* reading an anomaly (e.g., Braze et al., 2002; Ni et al., 1998; Rayner et al., 2004; Warren & McConnell, 2007). Moreover, increased reading times on the spillover region have been recently paralleled with increased N400 amplitude (Aurnhammer et al., 2021), which is modulated by violations at the level of semantics and reflect cognitive efforts in integrating a stimulus into a given context from semantic memory (Kliegl et al., 2012; Kutas & Federmeier, 2000; White et al., 2009). As such, the effects observed on the spillover may reflect semantic processing difficulties caused by stereotype-incongruent information. Consistent with this proposition, previous research with event-related brain potentials showed that stereotype-incongruent items elicited indeed larger N400 effects than stereotype-congruent items (Banaji & Hardin, 1996a; Bartholow & Dickter, 2008; White et al., 2009). Consequently, the present findings from the eye-tracking and self-paced reading studies, in line with research using event-related brain potentials, point to a close relationship between stereotype information and semantic memory[30].

Turning now to the effects of definite descriptions, the present study partially replicates previous findings with self-paced reading tasks (Mari & Müller, 2023, Chapter 4, this dissertation). In Experiment 1, first fixations on the second spillover were significantly longer when gender stereotype-incongruent information was presupposed through definite descriptions compared to asserted through indefinite descriptions. This finding is similar to

[30] Please note however that research with functional magnetic resonance imaging revealed that stereotypes also activate the same brain regions as when perceivers infer the beliefs, feelings, or opinions of others (Amodio, 2014; Contreras et al., 2012; Delplanque et al., 2019).

Mari and Müller's study, in which reading times were significantly slowed down on the spillover region when gender stereotype-incongruent information was presupposed. However, these effects were not observed for nationality stereotypes in Experiment 2 of the present study. This discrepancy in the results might be explained by two important considerations. First, the structure of the stimuli was different across the two experiments. Gender stereotype in/congruence occurred within the in/definite description (e.g., *la pilote*; the pilot$_{female}$), whereas nationality stereotype in/congruence stood in the relation between the in/definite description and the predicate (e.g., A/The French *seemed particularly polite*). Second, as previously presented, in/definite articles were most likely skipped in the present study, as functional words of 2-3 letters are often ignored during reading in natural settings. Consequently, eye-tracking methodology revealed to be inadequate for assessing subtle effects of linguistic markers of definiteness on information processing.

Overall, the present study revealed that when information violated gender stereotypes, and to a lesser extent nationality stereotypes, forward reading was interrupted, and attention shifted back to the locus of processing difficulty[31]. This finding suggests that readers' stereotypical expectations about gender occupations and inhabitants of certain countries influenced their online processing of the written text, but not its later processing. In future works, it will be important to directly examine the extent to which individual tendency to hold stereotypes affects the processing of information. Research using the Implicit Association Test (IAT; Greenwald et al., 1998), revealed that people display varying degrees of stereotype activation (Greenwald et al., 2003; Kurdi, Mann, et al., 2019; Nosek et al., 2007). In the present study, we did not assess individual differences in the extent to which participants showed automatic associations between category members (e.g., men, English people) and their properties (e.g., being a pilot, being polite). It is possible that people having high scores on the IAT (i.e., having strong stereotypical associations) show increased online and offline processing difficulties, but not people who obtained low scores on the IAT. In the future, research should take into account individual differences, and systematically assess how people's automatic stereotypical associations modulate information processing.

In sum, the present study provides a rigorous assessment of the specific processing time course of stereotype information during reading. The present results suggest that processing

[31] Please note that the findings of both studies do not allow us to draw firm conclusions about the research question as most of the analyses failed to reach statistical significance.

costs increased when reading a piece of information that violates stereotypical expectations about professional occupations and, to some extent, about inhabitants of different countries. Readers experienced difficulty integrating stereotype-incongruent information, interrupting their forward reading of the text and returning to the locus of processing difficulty. Put it simply, readers experienced the common "Wait, what?" reaction.

Summary and comments

Chapter 5 presented a further assessment with eye-tracking methodology of the specific processing time course of stereotype-incongruent information. Two experiments examined how reading was affected when information contradicted stereotypical expectations about gender and nationalities. The findings revealed that when information violated gender stereotypes, and to a lesser extent nationality stereotypes, forward reading was interrupted, and attention shifted back to the locus of processing difficulty. While this research offers some interesting insights into how and when stereotypical expectations get activated in natural reading settings, some important questions remain open.

First, most past studies assessed the effects gender stereotypes on reading, disregarding stereotypes about other social categories. In an endeavor to uncover the effects of stereotype activation during reading, future studies should seek to determine whether information processing is modulated by stereotypical representations about *any* social categories.

Second, like previous research in social psychology, the present study suffers from a sampling bias (Arnett, 2008), as it focused on student participants. This prevents us from drawing broad conclusions about stereotype activation. Therefore, future research should examine the extent to which contextual factors, cultural variations, personal motivations, or individual differences could mediate the activation of stereotypical expectations.

7. Conclusion

Collectively, Chapter 1 and Chapter 2 reveal that children readily make inferences about social categories, even when those categories are novel rather than existing. These chapters offer important insights into the early-emerging, possibly foundational, processes that underlie the acquisition of stereotypical expectations about social categories. Collectively, the studies reviewed in Chapter 1 support the theoretical perspective that language plays a crucial role in how children learn about social categories (Bigler & Liben, 2007; Bloom, 2004; Cimpian & Markman, 2011; Gelman, 2003; Over & McCall, 2018). When labels are used in combination with generic claims or high proportion quantifiers, children assume that any property applies to every member of a social category. However, if labels are combined with specific language, children interpret that the piece of information uniquely applies to one specific person. Chapter 1 thus confirms the perspective that language is determinant in social category learning, with labels signaling important and meaningful social divisions (Bigler & Liben, 2007; Over & McCall, 2018) and generic statements conveying information about the category *as a whole* (Bloom, 2004; Cimpian & Markman, 2011; Gelman, 2003). Chapter 1 also points to a need for research assessing how observed regularities shape children's representations of social categories. Determining whether children learn about social categories based on the regularities they observe in their environment represents a great opportunity to tackle the possible role of contextual factors in children's acquisition of stereotypes.

The experimental study presented in Chapter 2 addresses this latter question and complements previous research. The results support and extend the theoretical view that children can learn to associate properties with social categories from statistical evidence, i.e., by observing correlations between category members and certain properties (Bigler & Liben, 2006; Devine, 1989; S. J. Sherman et al., 2013; Shutts & Kalish, 2021). From 4 years old, children were able to generalize properties across members of the same novel social category based on their observation of regularities (i.e., after observing a majority of category members having the same property). Adding a novel perspective to this account, the study presented in Chapter 2 showed that starting from 7 years old, children are prone to make similarity inferences from their observations, without needing labels. Chapter 2 also addresses the theoretical perspective that the mere act of labeling social categories is interpreted by children as marking significant kinds of people (Bigler & Liben, 2007; Shutts & Kalish, 2021; Waxman & Markow,

1995). While 7-9-years old children generalized properties from simply learning about one labeled individual, younger children (4-6-year-olds) were more conservative in their generalization, attributing properties to social category members only after observing the *repeated* behaviors of verbally marked individuals. This dovetails nicely with previous research showing that young children might be more reluctant to draw inferences from small evidence, thereby preventing them from making overly broad generalizations in the process of early conceptual development (cf. Brandone et al., 2015). As children develop increased linguistic competence and experience school settings, they also become more sensitive to verbal cues to social categories and start to consider mere labels as meaningful markers of social divisions.

Whereas Chapter 1 and Chapter 2 explored how children come to attribute properties to others based on their category membership, Chapters 3 to 5 assessed how adult's stereotypical expectations about social categories modulate information processing during reading. Using reading tasks as an alternative method to capture the automatic activation of stereotypes, these studies enabled to probe stereotype activation in natural settings, without involving priming procedures or sorting tasks that might interrupt the ordinary flow of cognition. Two notable results emerged from this investigation. First, processing difficulties occurred immediately after reading items that contradicted stereotypical expectations. In contrast, reading was not affected when information was in line with stereotypical expectations. Second, information processing was affected by stereotypical expectations about gender and nationality to a similar extent. These findings converge on the conclusion that processing might be similarly affected by stereotypes about various social categories. Together these findings confirm theories of automatic stereotype activation, positing that information about social categories automatically increases the accessibility of stereotypic knowledge (D. E. Carlston, 1992; Fazio & Olson, 2003; Gilbert & Hixon, 1991; Greenwald & Banaji, 1995).

By using different methodologies, Chapter 4 and Chapter 5 offer complementary perspectives on stereotype activation during reading. First, the results from self-paced reading tasks in Chapter 4 indicate that reading slowed down immediately after the introduction of stereotype-incongruent information. Confirming theoretical accounts of stereotype activation during reading (Carreiras et al., 1996; Garnham et al., 2002; J. W. Sherman, Lee, et al., 1998; S. J. Sherman et al., 2013), Chapter 4 revealed that the mention of specific social category members in the text automatically activated stereotypical representations; when contradicted, these representations must be updated, resulting in increased reading times. Second, Chapter 5 employed an eye-tracking task, thereby offering a finer moment-by-moment analysis of how

the automatic activation of stereotypical expectations modulates information processing in natural reading settings. In line with theories of eye movement control during reading (Reichle, 2011; Reichle et al., 1998, 2003, 2009), the analysis of eye-movement patterns revealed that readers interrupted their forward reading of the text when encountering a piece of information that violated stereotypical expectations and returned to the locus of processing difficulty. Finally, in both studies reported in Chapter 4 and Chapter 5, stereotype-incongruent information led to increased reading times *after* reading the anomaly. From a theoretical point of view, this delayed effect is indicative of increased cognitive efforts to integrate a stimulus into a given context from semantic memory (Aurnhammer et al., 2021; Kliegl et al., 2012; Kutas & Federmeier, 2000; White et al., 2009). This finding speaks for the conceptual idea that stereotypes are organized into semantic networks, whereby social category information automatically activates associated concepts from semantic memory (D. E. Carlston, 1992; Devine, 1989; Fazio & Olson, 2003).

Altogether the present dissertation bears two important implications. First, the developmental part of this research revealed that children readily make inferences about others based on their category membership. While previous research focused on the role of language, and in particular labels and generic statements, the study presented in Chapter 2 showed that children of 7 years old learned to associate properties with social categories without verbal reference to the social category. This points to the potential critical role of children's environment and social experience in the acquisition of stereotypes. From early childhood, we are constantly exposed to a variety of contents that portray members of social categories in certain ways (e.g., men physicists, shy librarians, dangerous bikers, or punctual Swiss). Cumulatively, these social experiences might lead us to hold stereotypes about specific social categories.

The second major implication of this dissertation relates to the resulting effects of stereotypes. Upon reading information about social category members, a set of expectations about the likely attributes or behaviors of this social category immediately appears in people's minds. As revealed by the studies in Chapter 4 and Chapter 5, information that contradicted stereotypical expectations was costly to process. In contrast, when reading an item that matched stereotypical expectations, information processing was not affected. This points to the automatic nature of stereotypes, which function as mental shortcuts that generate expectations about social agents. Like heuristics, stereotypes facilitate information processing. This could in turn explain why, despite an increasing effort to promote equality and erase biased

representations of social groups in the 21st century, we still observed that information contradicting a stereotype is costly to process.

Focused on the cognitive underpinnings of stereotyping, the present dissertation shed light on the mechanisms by which we acquire and use stereotypical expectations about others. The present dissertation, however, set aside the social, affective, and contextual dimensions of stereotyping. Future research should seek to determine how these dimensions might interact in the formation and automatic activation of stereotypes. For instance, the question of the role of social learning in the formation of stereotypes remains open. While previous research showed that children's attitudes toward social categories are correlated with those of their parents (e.g., Degner & Dalege, 2013), there has been no direct empirical assessment of the way parents' – or peers' – communication about, and interactions with, members of a social category significantly affect the formation of stereotypes. Moreover, future research should also better integrate individual differences and cognitive processes involved in the formation and activation of stereotypes. How and to what extent would personal attitudes toward a social category, sense of group membership, or motivation to promote self-esteem systematically mediate the automatic activation of stereotypes? Together, these remaining questions represent very promising avenues for future research attempting to uncover the motives and affective factors that influence the acquisition and activation of stereotypes.

Overall, this dissertation advances our understanding of how we form stereotypical expectations about others and how they shape the way we process information. From an investigation of how children learn to associate properties with social categories, to the assessment of how stereotypical expectations modulate information processing, the present dissertation specified the conditions of stereotype acquisition and the cognitive processes involved in stereotyping. In addition, the present dissertation showed that we do form and hold stereotypical expectations about others, even in the absence of salient intergroup conflict or affective evaluations of social categories. This speaks to the view that stereotypes function as heuristics, which enable us to navigate the social world rapidly and efficiently; but stereotypes are a double-edged sword that can have serious consequences when applied wrongly, unjustly, or discriminatorily.

8. Appendixes

8.1. Appendix 1: List of stimuli and fillers used in Experiment 1 of Chapter 5

Condition: stereotype-congruent and indefinite NP

1. Yvette est allée à l'hôpital pour son diabète. Une aide soignante lui a été d'un grand soutien.
2. Eve est allée à la boucherie hier après-midi. Un boucher lui a vendu un rôti à cuire au four.
3. Lucienne est allée à l'hôpital le mois dernier. Un chirurgien l'a opérée avec une grande précision.
4. Esteban a appelé le service de dépannage électrique hier matin. Un électricien est venu chez lui tout de suite.
5. Mathieu est allé faire réparer sa voiture la semaine dernière. Un garagiste lui a prêté un véhicule pendant la durée des réparations.
6. Misha a contacté le service informatique hier soir. Un informaticien lui a répondu le lendemain.
7. Fabien est allé à une séance d'information sur l'armée hier soir. Un major lui a serré la main après son discours.
8. Jean a pris sa voiture pour aller au travail. Un motocycliste l'a dépassé à toute vitesse.
9. Pascal est allé contrôler les avions de Genève. Un pilote l'a regardé d'un air suspect.
10. Pénélope est allée au commissariat la semaine dernière. Un policier a pris sa déposition pour le vol de son portefeuille.
11. Aurore est allée en boîte vendredi dernier. Un videur lui a demandé de sortir.
12. Grégoire s'est fait cambrioler sa maison ce matin. Un voleur a tout pris en quelques heures.
13. Rose est allée au supermarché jeudi dernier. Une caissière lui a donné des bons de réductions.
14. Martine est allée chercher ses enfants à la garderie. Une éducatrice l'a saluée chaleureusement.
15. Léa est allée dans un salon de beauté lundi passé. Une esthéticienne l'a massée avec des pierres chaudes.
16. Serena est allée acheter des fleurs mercredi matin. Une fleuriste lui a conseillé de les mettre dans l'eau fraîche.
17. Robin est allé à l'hôpital mardi dernier. Une infirmière a mesuré sa tension artérielle.
18. Ludovic a regardé un documentaire sur la prostitution samedi soir. Une prostituée confiait vouloir changer de vie.
19. Paul est allé chez le docteur hier. Une secrétaire lui a fait une prise de sang.
20. César est allé dans un cabaret vendredi soir. Une serveuse lui a raconté des histoires coquines.
21. Dominique a fêté son divorce dans un cabaret. Une strip-teaseuse lui a offert une danse en privé.
22. Fabienne est allée consulter son avenir à Lausanne. Une voyante lui a prédit un avenir plein de succès.

Condition: stereotype-congruent and definite NP

23. Yvette est allée à l'hôpital pour son diabète. L'aide soignante lui a été d'un grand soutien.
24. Martine est allée chercher ses enfants à la garderie. L'éducatrice l'a saluée chaleureusement.
25. Esteban a appelé le service de dépannage électrique hier matin. L'électricien est venu chez lui tout de suite.
26. Léa est allée dans un salon de beauté lundi passé. L'esthéticienne l'a massée avec des pierres chaudes.
27. Robin est allé à l'hôpital mardi dernier. L'infirmière a mesuré sa tension artérielle.
28. Misha a contacté le service informatique hier soir. L'informaticien lui a répondu le lendemain.
29. Rose est allée au supermarché jeudi dernier. La caissière lui a donné des bons de réductions.
30. Serena est allée acheter des fleurs mercredi matin. La fleuriste lui a conseillé de les mettre dans l'eau fraîche.

31. Ludovic a regardé un documentaire sur la prostitution samedi soir. La prostituée confiait vouloir changer de vie.
32. Paul est allé chez le docteur hier. La secrétaire lui a fait une prise de sang.
33. César est allé dans un cabaret vendredi soir. La serveuse lui a raconté des histoires coquines.
34. Dominique a fêté son divorce dans un cabaret. La strip-teaseuse lui a offert une danse en privé.
35. Fabienne est allée consulter son avenir à Lausanne. La voyante lui a prédit un avenir plein de succès.
36. Eve est allée à la boucherie hier après-midi. Le boucher lui a vendu un rôti à cuire au four.
37. Lucienne est allée à l'hôpital le mois dernier. Le chirurgien l'a opérée avec une grande précision.
38. Mathieu est allé faire réparer sa voiture la semaine dernière. Le garagiste lui a prêté un véhicule pendant la durée des réparations.
39. Fabien est allé à une séance d'information sur l'armée hier soir. Le major lui a serré la main après son discours.
40. Jean a pris sa voiture pour aller au travail. Le motocycliste l'a dépassé à toute vitesse.
41. Pascal est allé contrôler les avions de Genève. Le pilote l'a regardé d'un air suspect.
42. Pénélope est allée au commissariat la semaine dernière. Le policier a pris sa déposition pour le vol de son portefeuille.
43. Aurore est allée en boîte vendredi dernier. Le videur lui a demandé de sortir.
44. Grégoire s'est fait cambrioler sa maison ce matin. Le voleur a tout pris en quelques heures.

Condition: stereotype-incongruent and definite NP

45. Yvette est allée à l'hôpital pour son diabète. L'aide soignant lui a été d'un grand soutien.
46. Martine est allée chercher ses enfants à la garderie. L'éducateur l'a saluée chaleureusement.
47. Esteban a appelé le service de dépannage électrique hier matin. L'électricienne est venue chez lui tout de suite.
48. Léa est allée dans un salon de beauté lundi passé. L'esthéticien l'a massée avec des pierres chaudes.
49. Robin est allé à l'hôpital mardi dernier. L'infirmier a mesuré sa tension artérielle.
50. Misha a contacté le service informatique hier soir. L'informaticienne lui a répondu le lendemain.
51. Eve est allée à la boucherie hier après-midi. La bouchère lui a vendu un rôti à cuire au four.
52. Lucienne est allée à l'hôpital le mois dernier. La chirurgienne l'a opérée avec une grande précision.
53. Mathieu est allé faire réparer sa voiture la semaine dernière. La garagiste lui a prêté un véhicule pendant la durée des réparations.
54. Fabien est allé à une séance d'information sur l'armée hier soir. La major lui a serré la main après son discours.
55. Jean a pris sa voiture pour aller au travail. La motocycliste l'a dépassé à toute vitesse.
56. Pascal est allé contrôler les avions de Genève. La pilote l'a regardé d'un air suspect.
57. Pénélope est allée au commissariat la semaine dernière. La policière a pris sa déposition pour le vol de son portefeuille.
58. Aurore est allée en boîte vendredi dernier. La videuse lui a demandé de sortir.
59. Grégoire s'est fait cambrioler sa maison ce matin. La voleuse a tout pris en quelques heures.
60. Rose est allée au supermarché jeudi dernier. Le caissier lui a donné des bons de réductions.
61. Serena est allée acheter des fleurs mercredi matin. Le fleuriste lui a conseilléde les mettre dans l'eau fraîche.
62. Ludovic a regardé un documentaire sur la prostitution samedi soir. Le prostitué confiait vouloir changer de vie.
63. Paul est allé chez le docteur hier. Le secrétaire lui a fait une prise de sang.
64. César est allé dans un cabaret vendredi soir. Le serveur lui a raconté des histoires coquines.
65. Dominique a fêté son divorce dans un cabaret. Le strip-teaseur lui a offert une danse en privé.
66. Fabienne est allée consulter son avenir à Lausanne. Le voyant lui a prédit un avenir plein de succès.

Condition: stereotype-incongruent and indefinite NP

67. Yvette est allée à l'hôpital pour son diabète. Un aide soignant lui a été d'un grand soutien.
68. Rose est allée au supermarché jeudi dernier. Un caissier lui a donné des bons de réductions.
69. Martine est allée chercher ses enfants à la garderie. Un éducateur l'a saluée chaleureusement.
70. Léa est allée dans un salon de beauté lundi passé. Un esthéticien l'a massée avec des pierres chaudes.
71. Serena est allée acheter des fleurs mercredi matin. Un fleuriste lui a conseillé de les mettre dans l'eau fraîche.
72. Robin est allé à l'hôpital mardi dernier. Un infirmier a mesuré sa tension artérielle.
73. Ludovic a regardé un documentaire sur la prostitution samedi soir. Un prostitué confiait vouloir changer de vie.
74. Paul est allé chez le docteur hier. Un secrétaire lui a fait une prise de sang.
75. César est allé dans un cabaret vendredi soir. Un serveur lui a raconté des histoires coquines.
76. Dominique a fêté son divorce dans un cabaret. Un strip-teaseur lui a offert une danse en privé.
77. Fabienne est allée consulter son avenir à Lausanne. Un voyant lui a prédit un avenir plein de succès.
78. Eve est allée à la boucherie hier après-midi. Une bouchère lui a vendu un rôti à cuire au four.
79. Lucienne est allée à l'hôpital le mois dernier. Une chirurgienne l'a opérée avec une grande précision.
80. Esteban a appelé le service de dépannage électrique hier matin. Une électricienne est venue chez lui tout de suite.
81. Mathieu est allé faire réparer sa voiture la semaine dernière. Une garagiste lui a prêté un véhicule pendant la durée des réparations.
82. Misha a contacté le service informatique hier soir. Une informaticienne lui a répondu le lendemain.
83. Fabien est allé à une séance d'information sur l'armée hier soir. Une major lui a serré la main après son discours.
84. Jean a pris sa voiture pour aller au travail. Une motocycliste l'a dépassé à toute vitesse.
85. Pascal est allé contrôler les avions de Genève. Une pilote l'a regardé d'un air suspect.
86. Pénélope est allée au commissariat la semaine dernière. Une policière a pris sa déposition pour le vol de son portefeuille.
87. Aurore est allée en boîte vendredi dernier. Une videuse lui a demandé de sortir.
88. Grégoire s'est fait cambrioler sa maison ce matin. Une voleuse a tout pris en quelques heures.

Fillers

1. Jeanne est allée à la gym ce matin. L'entraîneur lui parlait à voix haute.
2. John est allé à l'école lundi après-midi. L'enseignant suppléant a parlé avec lui brièvement.
3. Augustin est allé au spa durant le week-end. La masseuse l'a complimenté à propos de son apparence.
4. Mélanie est allée au bureau hier après le déjeuner. La secrétaire l'a aidée là-bas pendant une heure.
5. Kevin est allé au pub il y a trois jours. Le barman l'a reconnu tout suite.
6. Marie est allée dans un restaurant français cet automne. Le Chef lui a souhaité la bienvenue.
7. Fred est allé à un spectacle de danse le soir. Le chorégraphe lui a souri de loin.
8. Philippe est allé à la piscine mardi soir. Le coach de natation l'a insulté très grossièrement.
9. Susie est allée dans un salon de coiffure en automne. Le coiffeur s'est adressé à elle poliment.
10. Jeannette est allée au laboratoire l'après-midi. Le directeur du laboratoire l'a inspirée pour la science.
11. Ben est allé dans une clinique mercredi soir. Le docteur l'a salué calmement.
12. Lucien est allé en prison samedi soir. Le garde a parlé avec lui pendant un moment.
13. Loïc est allé au zoo lundi matin. Le gardien de zoo lui a crié dessus de manière inattendue.
14. Gabriella est allée à un concert il y a deux semaines. Le guitariste lui a fait un clin d'œil.
15. Anita est allée au cirque vendredi soir. Le lion a baillé devant elle plusieurs fois.
16. Louise est allée à la galerie d'art mardi soir. Le marchand d'art l'a intriguée avec un commentaire.
17. Mike est allé au port il y a deux jours Le pêcheur a parlé avec lui durant un long moment.
18. Dolorès est allée à la pharmacie vendredi après-midi. Le pharmacien l'a tracassée pendant dix minutes.

19. Rachelle est allée à la poste cet après-midi. Le postier l'a aidée de manière sympathique.
20. Harry est allé à la plage il y a quelques heures. Le sauveteur l'a averti du mauvais temps.
21. Catherine est allée dans un café ce matin. Le serveur l'a servie très rapidement.
22. Pedro est allé au tribunal hier après-midi. Le témoin l'a reconnu à travers la salle.
23. Marianne est allée au théâtre il y a deux jours Le vendeur de tickets a parlé avec elle durant un long moment.
24. Bill est allé dans un club vendredi soir. Le videur s'est disputé avec lui pendant un moment.

8.2. Appendix 2: Model selection procedure for Experiment 1 of Chapter 5

Online processing												
First fixation duration												
AOI1 (e.g., "A/The beautician")												
Model	**Model name**	**Model specification**	**Model fit**	**PCA**	**Model fit**				**Likelihood ratio test**			
					AIC	**BIC**	**LL**	**DF**	**Models compared**	**DF**	**X2**	***p*-value**
1	Maximal model (with log-transformed data)	lmer (log AOI1 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model failed to converge		1842.8	1969.1	-896.4	25				
2	Maximal model with bobyqa optimizer	lmer (log AOI1 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	1842.3	1968.7	-896.2	25	1 & 2	0	0.42	<.001
3	Model with zero-correlation-parameter (ZCP)	lmer (log AOI1 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	1824	1889.7	-899	13	2 & 3	12	5.62	0.934

Model	Model name	Model specification	Model fit	PCA	Model fit				Likelihood ratio test			
					AIC	BIC	LL	DF	Models compared	DF	X2	*p*-value
4	**Model ZCP dropping zero variance components**	**lmer (log AOI1 ~ stereotype * article + (stereotype || subjects) + (article || stimuli) Optimizer = bobyqa**			**1816**	**1861.5**	**-899**	**9**	**3 & 4**	**4**	**0**	**1**
5	Extending reduced model (4) with correlation	lmer (log AOI1 ~ stereotype * article + (stereotype | subjects) + (article | stimuli) Optimizer = bobyqa	Model failed to converge	Model overparametrized	1859.3	1914.9	-918.6	11	4 & 5	2	0	1
AOI2 - spillover (e.g., "offered her")												
1	Maximal model (with log-transformed data)	lmer (log AOI2 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model failed to converge		1568.4	1694.3	-759.2	25				
2	Maximal model with bobyqa optimizer	lmer (log AOI2 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	1568.4	1694.3	-759.2	25	1 & 2	0	0.00003	<.001
3	Model with zero-correlation-parameter (ZCP)	lmer (log AOI2 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	1548.4	1613.9	-761.2	13	2 & 3	12	4.05	0.982
4	**Model ZCP dropping zero variance components**	**lmer (log AOI2 ~ stereotype * article + (1 | subjects) + (article || stimuli) Optimizer = bobyqa**			**1538.4**	**1578.7**	**-761.2**	**8**	**3 & 4**	**5**	**0**	**1**

Model	Model name	Model specification	Model fit	PCA	Model fit				Likelihood ratio test			
					AIC	BIC	LL	DF	Models compared	DF	X2	*p*-value
5	Extending reduced model (4) with correlation	lmer (log AOI2 ~ stereotype * article + (1 | subjects) + (article | stimuli) Optimizer = bobyqa			1540.3	1585.7	-761.2	9	4 & 5	1	0.08	0.782
AOI3 - spillover (e.g., "a stone massage")												
1	Maximal model (log data)	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model failed to converge		1732.8	1859.2	-841.4	25				
2	Maximal model with nlminbwrap optimizer	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli) Optimizer: nlminbwrap	Model with singular fit	Model overparametrized	1732.8	1859.2	-841.4	25	1 & 2	0	0.01	<.001
3	Model zero-correlation-parameter (ZCP)	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli) Optimizer: nlminbwrap	Model with singular fit	Model overparametrized	1711.4	1777.1	-842.7	13	2 & 3	12	2.54	0.998
4	Model ZCP dropping zero variance components	lmer (log AOI3 ~ stereotype * article + (1 | subjects) + (0 + article || stimuli) Optimizer: nlminbwrap			1699.4	1734.7	-842.7	7	3 & 4	6	0	1
5	**Extending reduced model (4) with correlation**	**lmer (log AOI3 ~ stereotype * article + (1 | subjects) + (0 + article | stimuli) Optimizer: nlminbwrap**			**1699.4**	**1734.7**	**-842.7**	**7**	**4 & 5**	**0**	**0**	**1**

First-pass duration												
AOI1 (e.g., "A/The beautician")												
Model	**Model name**	**Model specification**	**Model fit**	**PCA**	**Model fit**				**Likelihood ratio test**			
					AIC	**BIC**	**LL**	**DF**	**Models compared**	**DF**	**X2**	***p*-value**
1	Maximal model (with log-transformed data)	lmer (log AOI1 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model failed to converge		1965.3	2091.7	-957.7	25				
2	Maximal model with bobyqa optimizer	lmer (log AOI1 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	1965.3	2091.7	-957.7	25	1 & 2	0	0.0003	<.001
3	Model with zero-correlation-parameter (ZCP)	lmer (log AOI1 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	1954.3	2020	-964.1	13	2 & 3	12	12.97	0.371
4	**Model ZCP dropping zero variance components**	**lmer (log AOI1 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype*article || stimuli) Optimizer = bobyqa**			**1950.3**	**2005.9**	**-964.1**	**11**	**3 & 4**	**2**	**0**	**1**
5	Extending reduced model (4) with correlation	lmer (log AOI1 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype*article | stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	1953	2044	-958.5	18	4 & 5	7	11.23	0.129

AOI2 - spillover (e.g., "offered her")												
Model	**Model name**	**Model specification**	**Model fit**	**PCA**	**Model fit**				**Likelihood ratio test**			
					AIC	**BIC**	**LL**	**DF**	**Models compared**	**DF**	**X2**	***p*-value**
1	Maximal model (with log-transformed data)	lmer (log AOI2 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model with singular fit	Model overparametrized	1787.9	1913.8	-868.9	25				
2	Model with zero-correlation-parameter (ZCP)	lmer (log AOI2 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli)	Model with singular fit	Model overparametrized	1774	1839.5	-874	13	1 & 2	12	10.13	0.6
3	**Model ZCP dropping zero variance components**	**lmer (log AOI2 ~ stereotype * article + (stereotype*article || subjects) + (article || stimuli)**			**1766**	**1811.3**	**-874**	**9**	**2 & 3**	**4**	**0**	**1**
4	Extending reduced model (3) with correlation	lmer (log AOI2 ~ stereotype * article + (stereotype*article | subjects) + (article | stimuli)	Model failed to converge		1766.9	1822.4	-872.5	11	3 & 4	2	3.06	0.217
AOI3 - spillover (e.g., "a stone massage")												
1	Maximal model (log data)	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model with singular fit	Model overparametrized	2557.8	2684.1	-1253.9	25				
2	Model zero-correlation-parameter (ZCP)	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli)	Model with singular fit	Model overparametrized	2541.6	2607.3	-1257.8	13	1 & 2	12	7.84	0.798

Model	Model name	Model specification	Model fit	PCA	Model fit				Likelihood ratio test			
					AIC	BIC	LL	DF	Models compared	DF	X2	*p*-value
3	Model ZCP dropping zero variance components	lmer (log AOI3 ~ stereotype * article + (stereotype + stereotype*article || subjects) + (stereotype + stereotype*article || stimuli)	Model with singular fit	Model overparametrized	2537.6	2593.2	-1257.8	11	2 & 3	2	0	1
4	**Model ZCP dropping variance components responsible for <1% of variance**	**lmer (log AOI3 ~ stereotype * article + (stereotype || subjects) + (stereotype || stimuli)**			**2533.6**	**2579.1**	**-1257.8**	**9**	**3 & 4**	**2**	**0.02**	**0.993**
5	Extending reduced model (4) with correlation	lmer (log AOI3 ~ stereotype * article + (stereotype | subjects) + (stereotype | stimuli)			2536	2591.6	-1257	11	4 & 5	2	1.58	0.454

Regression path time												
AOI1 (e.g., "A/The beautician")												
Model	**Model name**	**Model specification**	**Model fit**	**PCA**	**Model fit**				**Likelihood ratio test**			
					AIC	**BIC**	**LL**	**DF**	**Models compared**	**DF**	**X2**	***p*-value**
1	Maximal model (with log-transformed data)	lmer (log AOI1 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model with singular fit	Model overparametrized	2449.5	2575.8	-1199.7	25				
2	Model with zero-correlation-parameter (ZCP)	lmer (log AOI1 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli)	Model with singular fit	Model overparametrized	2439.9	2505.6	-1206.9	13	1 & 2	12	14.41	0.275
3	**Model ZCP dropping zero variance components**	**lmer (log AOI1 ~ stereotype * article + (stereotype + article || subjects) + (1 | stimuli)**			**2431.9**	**2477.4**	**-1206.9**	**9**	**2 & 3**	**4**	**0**	**1**
4	Extending reduced model (3) with correlation	lmer (log AOI1 ~ stereotype * article + (stereotype + article | subjects) + (1 | stimuli)	Model with singular fit	Model overparametrized	2433.8	2494.4	-1204.9	12	3 & 4	3	4.1	0.251
AOI2 - spillover (e.g., "offered her")												
1	Maximal model (with log-transformed data)	lmer (log AOI2 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model failed to converge		2724.7	2850.7	-1337.4	25				

Model	Model name	Model specification	Model fit	PCA	Model fit				Likelihood ratio test			
					AIC	BIC	LL	DF	Models compared	DF	X2	*p*-value
2	Maximal model with bobyqa optimizer	lmer (log AOI2 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	2703.3	2829.3	-1326.7	25	1 & 2	0	21.37	<.001
3	Model with zero-correlation-parameter (ZCP)	lmer (log AOI2 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	2693.2	2758.7	-1333.6	13	2 & 3	12	13.87	0.309
4	**Model ZCP dropping zero variance components**	**lmer (log AOI2 ~ stereotype * article + (1 | subjects) + (article || stimuli) Optimizer = bobyqa**			**2683.2**	**2723.5**	**-1333.6**	**8**	**3 & 4**	**5**	**0**	**1**
5	Extending reduced model (4) with correlation	lmer (log AOI2 ~ stereotype * article + (1 | subjects) + (article | stimuli) Optimizer = bobyqa			2682.8	2728.2	-1332.4	9	4 & 5	1	2.36	0.124
AOI3 - spillover (e.g., "a stone massage")												
1	Maximal model (log data)	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model failed to converge		2809.6	2935.9	-1379.8	25				
2	Maximal model with bobyqa optimizer	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	2807.2	2933.5	-1378.6	25	1 & 2	0	2.39	<.001

Model	Model name	Model specification	Model fit	PCA	Model fit				Likelihood ratio test			
					AIC	BIC	LL	DF	Models compared	DF	X2	*p*-value
3	Model zero-correlation-parameter (ZCP)	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	2799.9	2865.6	-1386.9	13	2 & 3	12	16.69	0.162
4	**Model ZCP dropping zero variance components**	**lmer (log AOI3 ~ stereotype * article + (stereotype + article || subjects) + (stereotype + article + stereotype*article || stimuli) Optimizer = bobyqa**			**2797.9**	**2858.5**	**-1386.9**	**12**	**3 & 4**	**1**	**0**	**1**
5	Extending reduced model (4) with correlation	lmer (log AOI3 ~ stereotype * article + (stereotype + article | subjects) + (stereotype + article + stereotype*article | stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	2803.9	2910	-1381	12	4 & 5	9	11.98	0.215

Offline processing

Total gaze duration

AOI1 (e.g., "A/The beautician")

Model	Model name	Model specification	Model fit	PCA	Model fit				Likelihood ratio test			
					AIC	BIC	LL	DF	Models compared	DF	X2	*p*-value
1	Maximal model (with log-transformed data)	lmer (log AOI1 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model with singular fit	Model overparametrized	2186.5	2312.9	-1068.3	25				
2	Model with zero-correlation-parameter (ZCP)	lmer (log AOI1 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli)	Model with singular fit	Model overparametrized	2179.2	2244.9	-1076.6	13	1 & 2	12	16.67	0.162
3	**Model ZCP dropping zero variance components**	**lmer (log AOI1 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (1 | stimuli)**			**2173.2**	**2223.8**	**-1076.6**	**10**	**2 & 3**	**3**	**0**	**1**
4	Extending reduced model (3) with correlation	lmer (log AOI1 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (1 | stimuli)	Model failed to converge		2174.9	2255.7	-1071.4	16	3 & 4	6	10.33	0.111

Sentence processing												
Model	**Model name**	**Model specification**	**Model fit**	**PCA**	**Model fit**				**Likelihood ratio test**			
					AIC	**BIC**	**LL**	**DF**	**Models compared**	**DF**	**X2**	***p*-value**
1	Maximal model (with log-transformed data)	lmer (log processing time ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model with singular fit	Model overparametrized	796.8	923.8	-373.4	25				
2	Model with zero-correlation-parameter (ZCP)	lmer (log processing time ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli)	Model with singular fit	Model overparametrized	792.3	858.3	-383.2	13	1 & 2	12	19.53	0.076
3	Model ZCP dropping zero variance components	**lmer (log processing time ~ stereotype * article + (article || subjects) + (stereotype*article || stimuli)**			784.3	830	-383.2	9	2 & 3	4	0	1
4	Extending reduced model (3) with correlation	lmer (log processing time ~ stereotype * article + (article | subjects) + (stereotype*article | stimuli)	Model failed to converge		787.9	843.7	-382.9	11	3 & 4	2	0.43	0.805

8.3. Appendix 3: List of stimuli and fillers used in Experiment 2 presented in Chapter 5

Condition: stereotype-congruent and indefinite NP

1. Antoine est allé en France l'été dernier. Un français lui a préparé des plats gastronomiques quelques soirs.
2. Manon est allée aux Etats-Unis le mois passé. Un américain lui a paru être vraiment obèse durant son séjour.
3. Nathan est allé en Angleterre l'hiver dernier. Un anglais lui a semblé particulièrement poli durant son voyage.
4. Céline est allée en Suède il y a deux semaines. Un suédois l'a sensibilisée à l'égalité des genres avec succès.
5. Manuel a conduit en Allemagne jeudi dernier. Un allemand respectait le code de la route de manière exemplaire.
6. Juliette est allée au Brésil il y a quatre mois. Un brésilien lui a appris l'art de danser la samba durant son séjour.
7. Ben est allé en Espagne le mois dernier. Un espagnol lui a paru être de nature bruyante la plupart du temps.
8. Marc est allé en Somalie pour une mission humanitaire. Un somalien lui a paru être vraiment très maigre durant sa mission.
9. Eric est allé en Russie il y a quelques jours. Un russes lui a révélé qu'il était homophobe sans aucune gêne.
10. Maria a étudié au Maroc durant les vacances d'été. Un marocain buvait du thé à la menthe toute la journée.
11. Jean a voyagé en Chine durant ses vacances. Un chinois lui a paru remarquablement petit comparé à lui.
12. Elodie est allée au Japon le week-end dernier. Un japonais lui a paru particulièrement timide envers les autres.
13. Marco est allé en Ecosse l'été dernier. Un écossais lui a semblé porter le kilt sans slip la plupart du temps.
14. Amélie est allée en Italie il y a deux mois. Un italien lui a présenté sa famille nombreuse de manière spontanée.
15. Patrick a conduit en Italie l'année passée. Un italien lui a semblé conduire comme un chauffard incorrigible.
16. Henry a vécu en Italie pendant une année. Un italien lui a parlé en bougeant les mains la plupart du temps.
17. Johanna est allée en Autriche suivre un cours de danse. Un autrichien lui a appris les mouvements de la valse avec beaucoup d'insistance.
18. Mathilde est allée en Italie le week-end dernier. Un italien a joué au grand séducteur tout le long du séjour.
19. Ophélie a étudié en Russie durant le semestre d'hiver. Un russe buvait beaucoup de vodka tous les soirs.
20. Paolo a voyagé au Kenya pendant ses vacances d'été. Un kényan lui a semblé être vraiment très pauvre comparé à lui.

Condition: stereotype-congruent and definite NP

21. Antoine est allé en France l'été dernier. Les français lui ont préparé des plats gastronomiques quelques soirs.
22. Manon est allée aux Etats-Unis le mois passé. Les américains lui ont paru être vraiment obèses durant son séjour.

23. Nathan est allé en Angleterre l'hiver dernier. Les anglais lui ont semblé particulièrement polis durant son voyage.
24. Céline est allée en Suède il y a deux semaines. Les suédois l'ont sensibilisée à l'égalité des genres avec succès.
25. Manuel a conduit en Allemagne jeudi dernier. Les allemands respectaient le code de la route de manière exemplaire.
26. Juliette est allée au Brésil il y a quatre mois. Les brésiliens lui ont appris l'art de danser la samba durant son séjour.
27. Ben est allé en Espagne le mois dernier. Les espagnols lui ont paru être de nature bruyante la plupart du temps.
28. Marc est allé en Somalie pour une mission humanitaire. Les somaliens lui ont paru être vraiment très maigres durant sa mission.
29. Eric est allé en Russie il y a quelques jours. Les russes lui ont révélé qu'ils étaient homophobes sans aucune gêne.
30. Maria a étudié au Maroc durant les vacances d'été. Les marocains buvaient du thé à la menthe toute la journée.
31. Jean a voyagé en Chine durant ses vacances. Les chinois lui ont paru remarquablement petits comparé à lui.
32. Elodie est allée au Japon le week-end dernier. Les japonais lui ont paru particulièrement timides envers les autres.
33. Marco est allé en Ecosse l'été dernier. Les écossais lui ont semblé porter le kilt sans slip la plupart du temps.
34. Amélie est allée en Italie il y a deux mois. Les italiens lui ont présenté leur famille nombreuse de manière spontanée.
35. Patrick a conduit en Italie l'année passée. Les italiens lui ont semblé conduire comme des chauffards incorrigibles.
36. Henry a vécu en Italie pendant une année. Les italiens lui ont parlé en bougeant les mains la plupart du temps.
37. Johanna est allée en Autriche suivre un cours de danse. Les autrichiens lui ont appris les mouvements de la valse avec beaucoup d'insistance.
38. Mathilde est allée en Italie le week-end dernier. Les italiens ont joué aux grands séducteurs tout le long du séjour.
39. Ophélie a étudié en Russie durant le semestre d'hiver. Les russes buvaient beaucoup de vodka tous les soirs.
40. Paolo a voyagé au Kenya pendant ses vacances d'été. Les kényans lui ont semblé être vraiment très pauvres comparé à lui.

Condition: stereotype-incongruent and definite NP

41. Antoine est allé en Angleterre l'été dernier. Les anglais lui ont préparé des plats gastronomiques quelques soirs.
42. Manon est allée en Thaïlande le mois passé. Les thaïlandais lui ont paru être vraiment obèses durant son séjour.
43. Nathan est allé en France l'hiver dernier. Les français lui ont semblé particulièrement polis durant son voyage.
44. Céline est allée en Arabie Saoudite il y a deux semaines. Les saoudiens l'ont sensibilisée à l'égalité des genres avec succès.
45. Manuel a conduit en Italie jeudi dernier. Les italiens respectaient le code de la route de manière exemplaire.

46. Juliette est allée en Pologne il y a quatre mois. Les polonais lui ont appris l'art de danser la samba durant son séjour.
47. Ben est allé au Japon le mois dernier. Les japonais lui ont paru être de nature bruyante la plupart du temps.
48. Marc est allé en Ecosse pour une mission humanitaire. Les écossais lui ont paru être vraiment très maigres durant sa mission.
49. Eric est allé au Canada il y a quelques jours. Les canadiens lui ont révélé qu'ils étaient homophobes sans aucune gêne.
50. Maria a étudié en Russie durant les vacances d'été. Les russes buvaient du thé à la menthe toute la journée.
51. Jean a voyagé en Norvège durant ses vacances. Les norvégiens lui ont paru remarquablement petits comparé à lui.
52. Elodie est allée en Espagne le week-end dernier. Les espagnols lui ont paru particulièrement timides envers les autres.
53. Marco est allé au Pakistan l'été dernier. Les pakistanais lui ont semblé porter le kilt sans slip la plupart du temps.
54. Amélie est allée en Suisse il y a deux mois. Les suisses lui ont présenté leur famille nombreuse de manière spontanée.
55. Patrick a conduit en Finlande l'année passée. Les finlandais lui ont semblé conduire comme des chauffards incorrigibles.
56. Henry a vécu en Allemagne pendant une année. Les allemands lui ont parlé en bougeant les mains la plupart du temps.
57. Johanna est allée à Cuba suivre un cours de danse. Les cubains lui ont appris les mouvements de la valse avec beaucoup d'insistance.
58. Mathilde est allée au Japon le week-end dernier. Les japonais ont joué aux grands séducteurs tout le long du séjour.
59. Ophélie a étudié en Algérie durant le semestre d'hiver. Les algériens buvaient beaucoup de vodka tous les soirs.
60. Paolo a voyagé au Qatar pendant ses vacances d'été. Les qatariens lui ont semblé être vraiment très pauvres. comparé à lui.

Condition: stereotype-incongruent and indefinite NP

61. Antoine est allé en Angleterre l'été dernier. Un anglais lui a préparé des plats gastronomiques quelques soirs.
62. Manon est allée en Thaïlande le mois passé. Un thaïlandais lui a paru être vraiment obèse durant son séjour.
63. Nathan est allé en France l'hiver dernier. Un français lui a semblé particulièrement poli durant son voyage.
64. Céline est allée en Arabie Saoudite il y a deux semaines. Un saoudien l'a sensibilisée à l'égalité des genres avec succès
65. Manuel a conduit en Italie jeudi dernier. Un italien respectait le code de la route de manière exemplaire.
66. Juliette est allée en Pologne il y a quatre mois. Un polonais lui a appris l'art de danser la samba durant son séjour.
67. Ben est allé au Japon le mois dernier. Un japonais lui a paru être de nature bruyante la plupart du temps.
68. Marc est allé en Ecosse pour une mission humanitaire. Un écossais lui a paru être vraiment très maigre durant sa mission.
69. Eric est allé au Canada il y a quelques jours. Un canadien lui a révélé qu'il était homophobe sans aucune gêne.
70. Maria a étudié en Russie durant les vacances d'été. Un russe buvait du thé à la menthe toute la journée.

71. Jean a voyagé en Norvège durant ses vacances. Un norvégien lui a paru remarquablement petit comparé à lui.
72. Elodie est allée en Espagne le week-end dernier. Un espagnol lui a paru particulièrement timide envers les autres.
73. Marco est allé au Pakistan l'été dernier. Un pakistanais lui a semblé porter le kilt sans slip la plupart du temps.
74. Amélie est allée en Suisse il y a deux mois. Un suisse lui a présenté sa famille nombreuse de manière spontanée.
75. Patrick a conduit en Finlande l'année passée. Un finlandais lui a semblé conduire comme un chauffard incorrigible.
76. Henry a vécu en Allemagne pendant une année. Un allemand lui a parlé en bougeant les mains la plupart du temps.
77. Johanna est allée à Cuba suivre un cours de danse. Un cubain lui a appris les mouvements de la valse avec beaucoup d'insistance.
78. Mathilde est allée au Japon le week-end dernier. Un japonais a joué au grand séducteur tout le long du séjour.
79. Ophélie a étudié en Algérie durant le semestre d'hiver. Un algérien buvait beaucoup de vodka tous les soirs.
80. Paolo a voyagé au Qatar pendant ses vacances d'été. Un qatarien lui a semblé être vraiment très pauvre comparé à lui.

Fillers

1. Gilles est allé en Grèce pendant ses vacances. Son hôtel se trouvait dans une réserve naturelle créée en 1980.
2. Veronica est allée au Brésil au mois de mars. Son séjour a été organisé par une agence locale.
3. Maxime est allé au Portugal l'été dernier. Son vol de retour a été annulé à plusieurs reprises à cause des intempéries.
4. Sabrina est allée au Maroc à Pacques. Ses valises ont été perdues par le personnel durant le voyage.
5. Barbara est allée au Brésil célébrer son mariage. Sa robe a été tachée dans sa valise durant le trajet.
6. Arthur est allé en Roumanie la semaine dernière. Son passeport arrivait à échéance le jour du retour.
7. Pablo est allé en Égypte cet été. Son billet d'avion était attitré à une autre voyageur du même vol.
8. Géraldine est allée aux Pays-Bas en train. Le wagon restaurant était fermé toute la durée du voyage.
9. Corentin est allé à Malte en bateau. Sa cabine était de la taille d'un cagibi.
10. Patricia est allée aux Maldives l'hiver dernier. Son bungalow se trouvait sur l'océan bleu turquoise.
11. Mario est allé en Chine le mois dernier. Son visa a été délivré par l'ambassade chinoise de Genève.
12. Jennifer est allée en Islande en avion. Son siège était situé à proximité d'une sortie de secours.
13. Alain est allé en Espagne l'hiver dernier. Son hôte l'a invité au musée le dernier jour.
14. Claire est allée aux Caraïbes pour son anniversaire. Ses amis lui ont offert un séjour dans un hôtel luxueux.
15. Ludovic est allé en Argentine pour un voyage d'affaire. Son collègue s'est perdu dans les rues de la capitale à plusieurs reprises.
16. Caroline est allée au Ghana en mission humanitaire. Son organisation était impliquée dans le soutien des enfants orphelins
17. Michel est allé en Éthiopie avec sa famille. Sa sœur a organisé un safari dans la savane
18. Gabriella est allée en Croatie faire des fouilles. Son équipe a retrouvé les vestiges d'une ancienne église.
19. Quentin est allé en Bulgarie pour les vacances. Son amie l'a rejoint quelques jours à la fin du séjour.
20. Méganne est allée en Malaisie pour le travail. Sa co-équipière l'a familiarisée avec les règlements internationaux.

21. Raphaël est allé au Canada avec son entreprise. Ses employés ont apprécié la ville de Québec et ses environs.
22. Victoria est allée au Chili pour des conférences. Son professeur lui a demandé de présenter son sujet de thèse.
23. Florent est allé au Viêt-Nam pour son mariage. Son témoin a organisé une cérémonie au bord de la mer.
24. Sylvie est allée en Finlande avec sa classe. Ses élèves ont apprécié visiter les musées de la capitale.

8.4. Appendix 4: Model selection procedure for Experiment 2 of Chapter 5

Online processing												
First fixation duration												
AOI1 (e.g., "A/The English / French")												
Model	**Model name**	**Model specification**	**Model fit**	**PCA**	**Model fit**				**Likelihood ratio test**			
					AIC	**BIC**	**LL**	**DF**	**Models compared**	**DF**	**X2**	***p*-value**
1	Maximal model (with log-transformed data)	lmer (log AOI1 ~ article + (article | subjects) + (article | stimuli)	Model failed to converge		1173.3	1215.8	-577.7	9				
2	Maximal model with bobyqa optimizer	lmer (log AOI1 ~ article + (article | subjects) + (article | stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	1160.5	1203	-571.3	9	1 & 2	0	12.84	<.001
3	Model with zero-correlation-parameter (ZCP)	lmer (log AOI1 ~ article + (article || subjects) + (article || stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	1158	1191.1	-572	7	2 & 3	2	1.54	0.463

Model	Model name	Model specification	Model fit	PCA	Model fit				Likelihood ratio test			
					AIC	BIC	LL	DF	Models compared	DF	X2	*p*-value
4	**Model ZCP dropping zero variance components**	**lmer (log AOI1 ~ article + (article || subjects) + (1 | stimuli) Optimizer = bobyqa**			**1156**	**1184.4**	**-572**	**6**	**3 & 4**	**1**	**0**	**1**
5	Extending reduced model (4) with correlation	lmer (log AOI1 ~ article + (article | subjects) + (1 | stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	1156.6	1189.7	-571.3	7	4 & 5	1	1.41	0.235
AOI2 - spillover (e.g., "seemed to be")												
1	Maximal model (with log-transformed data)	lmer (log AOI2 ~ article + (article | subjects) + (article | stimuli)	Model failed to converge		1162.8	1205.5	-572.4	9				
2	Maximal model with bobyqa optimizer	lmer (log AOI2 ~ article + (article | subjects) + (article | stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	1146.7	1189.4	-564.3	9	1 & 2	0	16.18	<.001
3	Model with zero-correlation-parameter (ZCP)	lmer (log AOI2 ~ article + (article || subjects) + (article || stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	1146.2	1179.4	-566.1	7	2 & 3	2	3.52	0.172
4	**Model ZCP dropping zero variance components**	**lmer (log AOI2 ~ article + (article || subjects) Optimizer = bobyqa**			**1142.2**	**1165.9**	**-566.1**	**5**	**3 & 4**	**2**	**0**	**1**
5	Extending model (4) with correlation	lmer (log AOI2 ~ article + (article | subjects) Optimizer = bobyqa	Model failed to converge	Model overparametrized	1158.1	1186.6	-573.1	6	4 & 5	1	0	1

AOI3 (e.g., "particularly polite")												
Model	**Model name**	**Model specification**	**Model fit**	**PCA**	**Model fit**				**Likelihood ratio test**			
					AIC	**BIC**	**LL**	**DF**	**Models compared**	**DF**	**X2**	***p*-value**
1	Maximal model (with log-transformed data)	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model failed to converge		1306.9	1426.3	-628.4	25				
2	Maximal model with bobyqa optimizer	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	1306.7	1426.2	-628.4	25	1 & 2	0	0.15	1
3	Model zero-correlation-parameter (ZCP)	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli)	Model with singular fit	Model overparametrized	1294.2	1356.3	-634.1	13	2 & 3	12	11.33	0.501
4	**Model ZCP dropping zero variance components**	**lmer (log AOI3 ~ stereotype * article + (article + stereotype*article || subjects) + (1 | stimuli)**			**1286.2**	**1329.2**	**-634.1**	**9**	**3 & 4**	**3**	**7.72**	**0.052**
5	Extending reduced model (4) with correlation	lmer (log AOI3 ~ stereotype * article + (article + stereotype*article | subjects) + (1 | stimuli)	Model with singular fit	Model overparametrized	1284.5	1341.8	-630.2	12	4 & 5	3	7.73	0.052

AOI4 - spillover (e.g., "during his stay.")												
Model	**Model name**	**Model specification**	**Model fit**	**PCA**	**Model fit**				**Likelihood ratio test**			
					AIC	**BIC**	**LL**	**DF**	**Models compared**	**DF**	**X2**	***p*-value**
1	Maximal model (with log-transformed data)	lmer (log AOI4 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model failed to converge		1406.2	1525.2	-678.1	25				
2	Maximal model with bobyqa optimizer	lmer (log AOI4 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	1406	1525.1	-678	25	1 & 2	0	0	1
3	Model zero-correlation-parameter (ZCP)	lmer (log AOI4 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli)	Model with singular fit	Model overparametrized	1391.9	1453.8	-682.9	13	2 & 3	12	9.23	0.631
4	**Model ZCP dropping zero variance components**	**lmer (log AOI4 ~ stereotype * article + (stereotype + stereotype*article || subjects) + (1 | stimuli)**			**1383.9**	**1426.7**	**-682.9**	**9**	**3 & 4**	**4**	**0.01**	**1**
5	Extending reduced model (4) with correlation	lmer (log AOI4 ~ stereotype * article + (article + stereotype*article | subjects) + (1 | stimuli)	Model with singular fit	Model overparametrized	1388.3	1445.4	-628.2	12	4 & 5	3	1.58	0.664

First-pass duration												
AOI1 (e.g., "A/The English / French")												
Model	**Model name**	**Model specification**	**Model fit**	**PCA**	**Model fit**				**Likelihood ratio test**			
					AIC	**BIC**	**LL**	**DF**	**Models compared**	**DF**	**X2**	***p*-value**
1	Maximal model (with log-transformed data)	lmer (log AOI1 ~ article + (article | subjects) + (article | stimuli)	Model failed to converge		1031.2	1073.7	-506.6	9				
2	Maximal model with bobyqa optimizer	lmer (log AOI1 ~ article + (article | subjects) + (article | stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	1031.2	1073.7	-506.6	9	1 & 2	0	0.015	1
3	Model with zero-correlation-parameter (ZCP)	lmer (log AOI1 ~ article + (article || subjects) + (article || stimuli) Optimizer = bobyqa	Model with singular fit	Model overparametrized	1029.8	1062.9	-507.9	7	2 & 3	2	2.61	0.271
4	**Model ZCP dropping zero variance components**	**lmer (log AOI1 ~ article + (article || subjects) + (1 | stimuli) Optimizer = bobyqa**			**1027.8**	**1056.2**	**-507.9**	**6**	**3 & 4**	**1**	**0**	**1**
5	Extending reduced model (4) with correlation	lmer (log AOI1 ~ article + (article | subjects) + (1 | stimuli) Optimizer = bobyqa			1029.2	1062.3	-507.6	7	4 & 5	1	0.58	0.445

AOI2 - spillover (e.g., "seemed to be")												
Model	**Model name**	**Model specification**	**Model fit**	**PCA**	**Model fit**				**Likelihood ratio test**			
					AIC	**BIC**	**LL**	**DF**	**Models compared**	**DF**	**X2**	***p*-value**
1	Maximal model (with log-transformed data)	lmer (log AOI2 ~ article + (article | subjects) + (article | stimuli)			1390	1432.7	-686	9				
AOI3 (e.g., "particularly polite")												
1	Maximal model (with log-transformed data)	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model with singular fit	Model overparametrized	2004.2	2123.7	-977.1	25				
2	Model zero-correlation-parameter (ZCP)	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli)	Model with singular fit	Model overparametrized	1989.1	2051.2	-981.5	13	1 & 2	12	8.82	0.718
3	**Model ZCP dropping zero variance components**	**lmer (log AOI3 ~ stereotype * article + (article || subjects) + (1 | stimuli)**			**1980**	**2018.3**	**-982**	**8**	**2 % 3**	**5**	**0.99**	**0.964**
4	Extending reduced model (3) with correlation	lmer (log AOI3 ~ stereotype * article + (article | subjects) + (1 | stimuli)			1982	2025.1	-982	9	3 & 4	1	0	0.993

AOI4 - spillover (e.g., "during his stay.")												
Model	**Model name**	**Model specification**	**Model fit**	**PCA**	**Model fit**				**Likelihood ratio test**			
					AIC	**BIC**	**LL**	**DF**	**Models compared**	**DF**	**X2**	***p*-value**
1	Maximal model (with log-transformed data)	lmer (log AOI4 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model with singular fit	Model overparametrized	1739.9	1858.9	-845	25				
2	Model zero-correlation-parameter (ZCP)	lmer (log AOI4 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli)	Model with singular fit	Model overparametrized	1729.1	1791	-851.6	13	1 & 2	12	13.21	0.354
3	**Model ZCP dropping zero variance components**	**lmer (log AOI4 ~ stereotype * article + (stereotype || subjects) + (stereotype*article || stimuli)**			**1721.1**	**1764**	**-851.6**	**9**	**2 & 3**	**4**	**0**	**1**
4	Extending reduced model (3) with correlation	lmer (log AOI4 ~ stereotype * article + (stereotype | subjects) + (stereotype*article | stimuli)		Model overparametrized	1722.6	1775	-850.3	11	3 & 4	2	2.49	0.287

Regression path time												
AOI3 (e.g., "particularly polite")												
1	Maximal model (with log-transformed data)	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model with singular fit	Model overparametrized	2138.8	2258.3	-1044.4	25				
2	Model zero-correlation-parameter (ZCP)	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli)	Model with singular fit	Model overparametrized	2121.8	2184	-1047.9	13	1 & 2	12	7.03	0.856
3	**Model ZCP dropping zero variance components**	**lmer (log AOI3 ~ stereotype * article + (article || subjects) + (1 | stimuli)**			**2111.8**	**2150.1**	**-1047.9**	**8**	**2 & 3**	**5**	**0**	**1**
4	Extending reduced model (3) with correlation	lmer (log AOI3 ~ stereotype * article + (article | subjects) + (1 | stimuli)			2113.8	2156.8	-1047.9	9	3 & 4	1	0.07	0.796

AOI4 - spillover (e.g., "during his stay.")												
Model	**Model name**	**Model specification**	**Model fit**	**PCA**	**Model fit**				**Likelihood ratio test**			
					AIC	**BIC**	**LL**	**DF**	**Models compared**	**DF**	**X2**	***p*-value**
1	Maximal model (with log-transformed data)	lmer (log AOI4 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model with singular fit	Model overparametrized	1946.2	2065.2	-948.1	25				
2	Model zero-correlation-parameter (ZCP)	lmer (log AOI4 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli)	Model with singular fit	Model overparametrized	1927.2	1989.1	-950.6	13	1 & 2	12	5.03	0.957
3	**Model ZCP dropping zero variance components**	**lmer (log AOI4 ~ stereotype * article + (stereotype || subjects) + (article || stimuli)**			**1919.2**	**1962.1**	**-950.6**	**9**	**2 & 3**	**4**	**0**	**1**
4	Extending reduced model (3) with correlation	lmer (log AOI4 ~ stereotype * article + (stereotype | subjects) + (article | stimuli)	Model failed to converge	Model overparametrized	1941.1	1993.5	-959.6	11	3 & 4	2	0	1

Offline processing

Total gaze duration

AOI1 (e.g., "A/The English / French")

Model	Model name	Model specification	Model fit	PCA	Model fit				Likelihood ratio test			
					AIC	BIC	LL	DF	Models compared	DF	X2	*p*-value
1	Maximal model (with log-transformed data)	lmer (log AOI1 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model with singular fit	Model overparametrized	1496	1614	-723	25				
2	Model with zero-correlation-parameter (ZCP)	lmer (log AOI1 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli)	Model with singular fit	Model overparametrized	1483.8	1545.2	-728.9	13	1 & 2	12	11.84	0.459
3	Model ZCP dropping zero variance components	lmer (log AOI1 ~ stereotype * article + (article + article*stereotype || subjects) + (stereotype + article || stimuli)	Model with singular fit	Model overparametrized	1479.8	1531.8	-728.9	11	2 & 3	2	0	1
4	**Model ZCP dropping variance components responsible for <1% of variance**	**lmer (log AOI1 ~ stereotype * article + (article || subjects) + (article || stimuli)**			**1475.8**	**1518.3**	**-728.9**	**9**	**3 & 4**	**2**	**0**	**1**
5	Extending reduced model (4) with correlation	lmer (log AOI1 ~ stereotype * article + (article | subjects) + (article | stimuli)	Model with singular fit	Model overparametrized	1476.4	1528.3	-727.2	11	4 & 5	2	3.43	0.18

AOI3 (e.g., "particularly polite")												
Model	**Model name**	**Model specification**	**Model fit**	**PCA**	**Model fit**				**Likelihood ratio test**			
					AIC	**BIC**	**LL**	**DF**	**Models compared**	**DF**	**X2**	***p*-value**
1	Maximal model (with log-transformed data)	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model with singular fit	Model overparametrized	1531.7	1651.2	-740.8	25				
2	Model zero-correlation-parameter (ZCP)	lmer (log AOI3 ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli)	Model with singular fit	Model overparametrized	1509.7	1571.8	-741.9	13	1 & 2	12	2.04	0.999
3	Model ZCP dropping zero variance components	lmer (log AOI3 ~ stereotype * article + (article + stereotype*article || subjects) + (stereotype + article || stimuli)	Model with singular fit		1505.7	1558.3	-741.9	11	2 & 3	2	0	1
4	**Model ZCP dropping variance components responsible for <1% of variance**	**lmer (log AOI3 ~ stereotype * article + (article || subjects) + (article || stimuli)**			**1501.7**	**1544.7**	**-741.9**	**9**	**3 & 4**	**2**	**0**	**1**
5	Extending reduced model (4) with correlation	lmer (log AOI3 ~ stereotype * article + (article | subjects) + (article | stimuli)			1505.2	1557.8	-741.6	11	4 & 5	2	0.48	0.786

Sentence processing												
Model	**Model name**	**Model specification**	**Model fit**	**PCA**	**Model fit**				**Likelihood ratio test**			
					AIC	**BIC**	**LL**	**DF**	**Models compared**	**DF**	**X2**	***p*-value**
1	Maximal model (with log-transformed data)	lmer (log processing time ~ stereotype * article + (stereotype + article + stereotype*article | subjects) + (stereotype + article + stereotype*article | stimuli)	Model with singular fit	Model overparametrized	496.9	616.9	-223.5	25				
2	Model with zero-correlation-parameter (ZCP)	lmer (log processing time ~ stereotype * article + (stereotype + article + stereotype*article || subjects) + (stereotype + article + stereotype*article || stimuli)	Model with singular fit	Model overparametrized	480	542.4	-227	13	1 & 2	12	7.11	0.85
3	**Model ZCP dropping zero variance components**	**lmer (log processing time ~ stereotype * article + (article || subjects) + (stereotype + article || stimuli)**			**474**	**522**	**-227**	**10**	**2 & 3**	**3**	**0**	**1**
4	Extending reduced model (3) with correlation	lmer (log processing time ~ stereotype * article + (article | subjects) + (stereotype + article | stimuli)	Model failed to converge	Model overparametrized	479.8	547.1	-225.9	14	3 & 4	4	2.18	0.702

9. References

Aboud, F. (1988). *Children and prejudice*. Blackwell Publishers.

Acorn, D. A., Hamilton, D. L., & Sherman, S. J. (1988). Generalization of Biased Perceptions of Groups Based on Illusory Correlations. *Social Cognition*, *6*(4), 345–372. https://doi.org/10.1521/soco.1988.6.4.345

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

Aloise, P. A. (1993). Trait confirmation and disconfirmation: The development of attribution biases. *Journal of Experimental Child Psychology*, *55*(2), 177–193. https://doi.org/10.1006/jecp.1993.1010

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

Alves, H., Koch, A., & Unkelbach, C. (2017b). Why Good Is More Alike Than Bad: Processing Implications. *Trends in Cognitive Sciences*, *21*(2), 69–79. https://doi.org/10.1016/j.tics.2016.12.006

Alves, H., Koch, A., & Unkelbach, C. (2018). A Cognitive-Ecological Explanation of Intergroup Biases. *Psychological Science*, *29*(7), 1126–1133. https://doi.org/10.1177/0956797618756862

Amodio, D. M. (2014). The neuroscience of prejudice and stereotyping. *Nature Reviews Neuroscience*, *15*(10), 670–682. https://doi.org/10.1038/nrn3800

Anderson, J. R., & Bower, G. H. (1973). *Human Associative Memory*. Winston. https://doi.org/10.4324/9781315802886

Anzures, G., Quinn, P. C., Pascalis, O., Slater, A. M., & Lee, K. (2009). Categorization, categorical perception, and asymmetry in infants' representation of face race. *Developmental Science*, *13*(4), 553–564. https://doi.org/10.1111/j.1467-7687.2009.00900.x

Arnett, J. J. (2008). The neglected 95%: Why American psychology needs to become less American. *American Psychologist*, *63*(7), 602–614. https://doi.org/10.1037/0003-066X.63.7.602

Ashburn-Nardo, L., & Johnson, N. J. (2008). Implicit Outgroup Favoritism and Intergroup Judgment: The Moderating Role of Stereotypic Context. *Social Justice Research*, *21*(4), 490–508. https://doi.org/10.1007/s11211-008-0078-8

Aurnhammer, C., Delogu, F., Schulz, M., Brouwer, H., & Crocker, M. W. (2021). Retrieval (N400) and integration (P600) in expectation-based comprehension. *PLOS ONE*, *16*(9), e0257430. https://doi.org/10.1371/journal.pone.0257430

Banaji, M. R., & Greenwald, A. G. (2013). Blindspot: Hidden biases of good people. In *Blindspot: Hidden biases of good people.* Delacorte Press.

Banaji, M. R., & Hardin, C. D. (1996a). Automatic stereotyping. *Psychological Science*, *7*(3), 136–141. https://doi.org/10.1111/j.1467-9280.1996.tb00346.x

Banaji, M. R., & Hardin, C. D. (1996b). Automatic stereotyping. *Psychological Science*, *7*(3), 136–141. https://doi.org/10.1111/j.1467-9280.1996.tb00346.x

Bargh, J. A., & Pietromonaco, P. (1982). Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology*, *43*(3), 437–449. https://doi.org/10.1037/0022-3514.43.3.437

Baron, A. S. (2009). *Foundations of intergroup cognition* [Dissertation, Harvard University]. Available from ProQuest Dissertations & Theses Global database. (UMI No. 3395407)

Baron, A. S., & Dunham, Y. (2015). Representing 'us' and 'them': Building blocks of intergroup cognition. *Journal of Cognition and Development*, *16*(5), 780–801. https://doi.org/10.1080/15248372.2014.1000459

Baron, A. S., Dunham, Y., Banaji, M., & Carey, S. (2014). Constraints on the acquisition of social category concepts. *Journal of Cognition and Development*, *15*(2), 238–268. https://doi.org/10.1080/15248372.2012.742902

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bar-Tal, D. (1996). Development of social categories and stereotypes in early childhood: The case of "The Arab" concept formation, stereotype and attitudes by Jewish children in Israel. *International Journal of Intercultural Relations*, *20*(3–4), 341–370. https://doi.org/10.1016/0147-1767(96)00023-5

Bartholow, B. D., & Dickter, C. L. (2008). A response conflict account of the effects of stereotypes on racial categorization. *Social Cognition*, *26*(3), 314–332. https://doi.org/10.1521/soco.2008.26.3.314

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *ArXiv:1506.04967*, 1–21.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–51. https://doi.org/10.18637/jss.v067.i01

Benozio, A., & Diesendruck, G. (2015). Parochialism in preschool boys' resource allocation. *Evolution and Human Behavior*, *36*(4), 256–264. https://doi.org/10.1016/j.evolhumbehav.2014.12.002

Beukeboom, C. J., & Burgers, C. (2019). How stereotypes are shared through language: A review and introduction of the social categories and stereotypes communication (SCSC) framework. *Review of Communication Research*, *7*, 1–37. https://doi.org/10.12840/issn.2255-4165.017

Bigler, R. S., & Liben, L. S. (2006). A developmental intergroup theory of social stereotypes and prejudice. *Advances in Child Development and Behavior*, *34*, 39–89. https://doi.org/10.1016/s0065-2407(06)80004-2

Bigler, R. S., & Liben, L. S. (2007). Developmental Intergroup Theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science*, *16*(3), 162–166. https://doi.org/10.1111/j.1467-8721.2007.00496.x

Bigler, R. S., Spears Brown, C., & Markell, M. (2001). When groups are not created equal: Effects of group status on the formation of intergroup attitudes in children. *Child Development*, *72*(4), 1151–1162. https://doi.org/10.1111/1467-8624.00339

Blair, I. v. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, *6*(3), 242–261. https://doi.org/10.1207/S15327957PSPR0603_8

Blanchard, H. E., & Iran-Nejad, A. (1987). Comprehension processes and eye movement patterns in the reading of surprise-ending stories. *Discourse Processes*, *10*(1), 127–138. https://doi.org/10.1080/01638538709544663

Bloom, P. (2004). *Descartes' baby: How the science of child development explains what makes us human*. Basic Books.

Boseovski, J. J., & Lee, K. (2006). Children's use of frequency information for trait categorization and behavioral prediction. *Developmental Psychology*, *42*(3), 500–513. https://doi.org/10.1037/0012-1649.42.3.500

Bourne, L. E. (1966). *Human conceptual behavior*. Allyn and Bacon.

Brandone, A. C., Gelman, S. A., & Hedglen, J. (2015). Children's developing intuitions about the truth conditions and implications of novel generics versus quantified statements. *Cognitive Science*, *39*(4), 711–738. https://doi.org/10.1111/cogs.12176

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, *23*(3), 389–411. https://doi.org/10.1037/met0000159

Braze, D., Shankweiler, D., Ni, W., & Palumbo, L. C. (2002). Readers' eye movements distinguish anomalies of form and content . *Journal of Psycholinguistic Research*, *31*(1), 25–44. https://doi.org/10.1023/A:1014324220455

Brewer, M. B. (1988). A dual process model of impression formation. In R. S. Wyer & T. K. Srull (Eds.), *Advances in Social Cognition* (Vol. 1, pp. 1–36). Lawrence Erlbaum Associates.

Brey, E., & Pauker, K. (2019). Teachers' nonverbal behaviors influence children's stereotypic beliefs. *Journal of Experimental Child Psychology*, *188*, 104671. https://doi.org/10.1016/j.jecp.2019.104671

Brey, E., & Shutts, K. (2018). Children use nonverbal cues from an adult to evaluate peers. *Journal of Cognition and Development*, *19*(2), 121–136. https://doi.org/10.1080/15248372.2018.1449749

Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. Lloyd (Eds.), *Cognition und calegorization* (pp. 169–211). Lawrence Erlbaum.

Brown, C. S., & Bigler, R. S. (2002). Effects of minority status in the classroom on children's intergroup attitudes. *Journal of Experimental Child Psychology*, *83*(2), 77–110. https://doi.org/10.1016/S0022-0965(02)00123-6

Brown, P. M., & Turner, J. C. (2002). The role of theories in the formation of stereotype content. In C. McGarty, V. Y. Yzerbyt, & R. Spears (Eds.), *Stereotypes as explanations: The formation of meaningful beliefs about social groups* (pp. 67–89). Cambridge University Press. https://doi.org/10.1017/CBO9780511489877.005

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

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. John Wiley and Sons.

Buchsbaum, D., Gopnik, A., Griffiths, T. L., & Shafto, P. (2011). Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence. *Cognition*, *120*(3), 331–340. https://doi.org/10.1016/j.cognition.2010.12.001

Cacciari, C., Carreiras, M., & Cionini, C. B. (1997). When words have two genders: Anaphor resolution for Italian functionally ambiguous words. *Journal of Memory and Language*, *37*(4), 517–532. https://doi.org/10.1006/jmla.1997.2528

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., … Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*, *2*(9), 637–644. https://doi.org/10.1038/s41562-018-0399-z

Campbell, D. T. (1967). Stereotypes and the perception of group differences. *American Psychologist*, *22*(10), 817–829. https://doi.org/10.1037/h0025079

Cantor, N., & Mischel, W. (1979). Prototypes in Person Perception. *Advances in Experimental Social Psychology*, *12*, 3–52. https://doi.org/10.1016/S0065-2601(08)60258-0

Carlston, D. (2010). Models of implicit and explicit mental representation. In B. Gawronski & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, theory, and applications* (pp. 38–79). Guilford Press.

Carlston, D. E. (1992). Impression formation and the modular mind: The associated systems theory. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 301–341). Lawrence Erlbaum Associates.

Carpenter, P. A., & Just, M. A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 275–307). Academic Press. https://doi.org/10.1016/B978-0-12-583680-7.50022-9

Carreiras, M., Garnham, A., Oakhill, J., & Cain, K. (1996). The use of stereotypical gender information in constructing a mental model: Evidence from English and Spanish. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *49*(3), 639–663. https://doi.org/10.1080/027249896392531

Cave, T., & Wilson, D. (2018). *Reading Beyond the Code: Literature and Relevance Theory*. Oxford University Press.

Cesario, J. (2014). Priming, replication, and the hardest science. *Perspectives on Psychological Science*, *9*(1), 40–48. https://doi.org/10.1177/1745691613513470

Chalik, L., & Dunham, Y. (2020). Beliefs about moral obligation structure children's social category-based expectations. *Child Development*, *91*(1), e108–e119. https://doi.org/10.1111/cdev.13165

Chalik, L., & Rhodes, M. (2014). Preschoolers use social allegiances to predict behavior. *Journal of Cognition and Development*, *15*(1), 136–160. https://doi.org/10.1080/15248372.2012.728546

Chalik, L., & Rhodes, M. (2015). The communication of naïve theories of the social world in parent-child conversation. *Journal of Cognition and Development*, *16*(5), 719–741. https://doi.org/10.1080/15248372.2014.949722

Chalik, L., & Rhodes, M. (2018). Learning about social category-based obligations. *Cognitive Development*, *48*, 117–124. https://doi.org/10.1016/j.cogdev.2018.06.010

Chalik, L., Rivera, C., & Rhodes, M. (2014). Children's use of categories and mental states to predict social behavior. *Developmental Psychology*, *50*(10), 2360–2367. https://doi.org/10.1037/a0037729

Chambers, C. G., Graham, S. A., & Turner, J. N. (2008). When hearsay trumps evidence: How generic language guides preschoolers' inferences about unfamiliar things. *Language and Cognitive Processes*, *23*(5), 749–766. https://doi.org/10.1080/01690960701786111

Charlesworth, T. E. S., & Banaji, M. R. (n.d.). The development of social group cognition: What infants and young children can teach us. In *Oxford Handbook of Social Cognition* (2nd Edition). Retrieved May 30, 2022, from http://www.people.fas.harvard.edu/~banaji/research/publications/articles/2021_Charlesworth_OHSC.pdf

Charlesworth, T. E. S., & Banaji, M. R. (2019). Patterns of Implicit and Explicit Attitudes: I. Long-Term Change and Stability From 2007 to 2016. *Psychological Science*, *30*(2), 174–192. https://doi.org/10.1177/0956797618813087

Charlesworth, T. E. S., Kurdi, B., & Banaji, M. R. (2020). Children's implicit attitude acquisition: Evaluative statements succeed, repeated pairings fail. *Developmental Science*, *23*(3), e12911. https://doi.org/10.1111/desc.12911

Charlesworth, T. E. S., Yang, V., Mann, T. C., Kurdi, B., & Banaji, M. R. (2021). Gender stereotypes in natural language: Word embeddings show robust consistency across child and adult language corpora of more than 65 million words. *Psychological Science*, *32*(2), 218–240. https://doi.org/10.1177/0956797620963619

Chudek, M., & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences*, *15*(5), 218–226. https://doi.org/10.1016/j.tics.2011.03.003

Cimpian, A., & Markman, E. M. (2011). The generic/nongeneric distinction influences how children interpret new information about social others. *Child Development*, *82*(2), 471–492. https://doi.org/10.1111/j.1467-8624.2010.01525.x

Cimpian, A., & Salomon, E. (2014). The inherence heuristic: An intuitive means of making sense of the world, and a potential precursor to psychological essentialism. *Behavioral and Brain Sciences*, *37*(5), 461–480. https://doi.org/10.1017/S0140525X13002197

Clark, H. H. (1975). Bridging. *Proceedings of the 1975 Workshop on Theoretical Issues in Natural Language Processing - TINLAP '75*, 169–174. https://doi.org/10.3115/980190.980237

Clément, F. (2010). De la nature des croyances collectives. *L'Année Sociologique*, *60*(1), 63–91. https://doi.org/10.3917/anso.101.0063

Clément, F., & Dukes, D. (2017). Social appraisal and social referencing: Two components of affective social learning. *Emotion Review*, *9*(3), 253–261. https://doi.org/10.1177/1754073916661634

Clifton, C. (2013). Situational context affects definiteness preferences: Accommodation of presuppositions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 487–501. https://doi.org/10.1037/a0028975

Cohen, J. B., & Basu, K. (1987). Alternative Models of Categorization: Toward a Contingent Processing Framework. *Journal of Consumer Research*, *13*(4), 455. https://doi.org/10.1086/209081

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

Contreras, J. M., Banaji, M. R., & Mitchell, J. P. (2012). Dissociable neural correlates of stereotypes and other forms of semantic knowledge. *Social Cognitive and Affective Neuroscience*, *7*(7), 764–770. https://doi.org/10.1093/scan/nsr053

Coronel, J. C., & Federmeier, K. D. (2016). The Effects of Gender Cues and Political Sophistication on Candidate Evaluation. *Communication Research*, *43*(7), 922–944. https://doi.org/10.1177/0093650215604024

Crosby, F., Bromley, S., & Saxe, L. (1980). Recent unobtrusive studies of Black and White discrimination and prejudice: A literature review. *Psychological Bulletin*, *87*(3), 546–563. https://doi.org/10.1037/0033-2909.87.3.546

Cuddy, A. J. C., Fiske, S. T., Kwan, V. S. Y., Glick, P., Demoulin, S., Leyens, J.-P., Bond, M. H., Croizet, J.-C., Ellemers, N., Sleebos, E., Htun, T. T., Kim, H.-J., Maio, G., Perry, J., Petkova, K., Todorov, V., Rodríguez-Bailón, R., Morales, E., Moya, M., … Ziegler, R. (2009). Stereotype content model across cultures: Towards universal similarities and some differences. *British Journal of Social Psychology*, *48*(1), 1–33. https://doi.org/10.1348/014466608X314935

Cvencek, D., Meltzoff, A. N., Maddox, C. D., Nosek, B. A., Rudman, L. A., Devos, T., Dunham, Y., Baron, A. S., Steffens, M. C., Lane, K., Horcajo, J., Ashburn-Nardo, L., Quinby, A., Srivastava, S. B., Schmidt, K., Aidman, E., Tang, E., Farnham, S., Mellott, D. S., … Greenwald, A. G. (2021). Meta-Analytic Use of Balanced Identity Theory to Validate the Implicit Association Test. *Personality and Social Psychology Bulletin*, *47*(2), 185–200. https://doi.org/10.1177/0146167220916631

Daneman, M., Reingold, E. M., & Davidson, M. (1995). Time course of phonological activation during reading: Evidence from eye fixations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 884–898. https://doi.org/10.1037/0278-7393.21.4.884

Davidson, N. S., & Gelman, S. A. (1990). Inductions from novel categories: The role of language and conceptual structure. *Cognitive Development*, *5*(2), 151–176. https://doi.org/10.1016/0885-2014(90)90024-N

Davis-Kean, P. E., & Ellis, A. (2019). An overview of issues in infant and developmental research for the creation of robust and replicable science. *Infant Behavior and Development*, *57*, 101339. https://doi.org/10.1016/j.infbeh.2019.101339

de Houwer, J., & Moors, A. (2010). Implicit measures: Similarities and differences. In B. Gawronski & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, theory, and applications* (pp. 176–193). Guilford Press.

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

de Rosnay, M., Cooper, P. J., Tsigaras, N., & Murray, L. (2006). Transmission of social anxiety from mother to infant: An experimental study using a social referencing paradigm. *Behaviour Research and Therapy*, *44*(8), 1165–1175. https://doi.org/10.1016/j.brat.2005.09.003

Degner, J., & Dalege, J. (2013). The apple does not fall far from the tree, or does it? A meta-analysis of parent–child similarity in intergroup attitudes. *Psychological Bulletin*, *139*(6), 1270–1304. https://doi.org/10.1037/a0031436

Delplanque, J., Heleven, E., & van Overwalle, F. (2019). Neural representations of groups and stereotypes using fMRI repetition suppression. *Scientific Reports*, *9*(1), 3190. https://doi.org/10.1038/s41598-019-39859-y

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

Dickinson, J. (2011). The impact of 'violating the heterosexual norm' on reading speed and accuracy. *Psychology*, *02*(05), 456–459. https://doi.org/10.4236/psych.2011.25071

Diesendruck, G. (2020). Why do children essentialize social groups? In *Advances in Child Development and Behavior* (Vol. 59, pp. 31–64). Elsevier. https://doi.org/10.1016/bs.acdb.2020.05.002

Diesendruck, G. (2021). A motivational perspective on the development of social essentialism. *Current Directions in Psychological Science*, *30*(1), 76–81. https://doi.org/10.1177/0963721420980724

Diesendruck, G., & Deblinger-Tangi, R. (2014). The linguistic construction of social categories in toddlers. *Child Development*, *85*(1), 114–123. https://doi.org/10.1111/cdev.12130

Diesendruck, G., & Eldror, E. (2011). What children infer from social categories. *Cognitive Development*, *26*(2), 118–126. https://doi.org/10.1016/j.cogdev.2010.11.001

Diesendruck, G., Gelman, S. A., & Lebowitz, K. (1998). Conceptual and linguistic biases in children's word learning. *Developmental Psychology*, *34*(5), 823–839. https://doi.org/10.1037/0012-1649.34.5.823

Diesendruck, G., & haLevi, H. (2006). The role of language, appearance, and culture in children's social category-based induction. *Child Development*, *77*(3), 539–553. https://doi.org/10.1111/j.1467-8624.2006.00889.x

Diesendruck, G., & Peretz, S. (2013). Domain differences in the weights of perceptual and conceptual information in children's categorization. *Developmental Psychology*, *49*(12), 2383–2395. https://doi.org/10.1037/a0032049

Diesendruck, G., Salzer, S., Kushnir, T., & Xu, F. (2015). When choices are not personal: The effect of statistical and social cues on children's inferences about the scope of preferences. *Journal of Cognition and Development*, *16*(2), 370–380. https://doi.org/10.1080/15248372.2013.848870

Diesendruck, G., & Weiss, E. (2015). Children's differential weighting of cues to social categories. *Cognitive Development*, *33*, 56–72. https://doi.org/10.1016/j.cogdev.2014.06.001

Doan, T., Friedman, O., & Denison, S. (2021). Toddlers and preschoolers understand that some preferences are more subjective than others. *Child Development*, *92*(3), 853–861. https://doi.org/10.1111/cdev.13581

Dovidio, J. F. (1999). Stereotyping. In R. A. Wilson & F. C. Keil (Eds.), *The MIT Encyclopedia of the Cognitive Science* (pp. 804–806). MIT Press. https://www.proquest.com/books/mit-encyclopedia-cognitive-sciences/docview/619413801/se-2?accountid=14610

Dovidio, J. F., Evans, N., & Tyler, R. B. (1986). Racial stereotypes: The contents of their cognitive representations. *Journal of Experimental Social Psychology*, *22*(1), 22–37. https://doi.org/10.1016/0022-1031(86)90039-9

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, *82*(1), 62–68. https://doi.org/10.1037/0022-3514.82.1.62

Duffy, S. A., & Keir, J. A. (2004). Violating stereotypes: Eye movements and comprehension processes when text conflicts with world knowledge. *Memory & Cognition*, *32*(4), 551–559. https://doi.org/10.3758/BF03195846

Dunbar, R. I. M. (2004). Gossip in evolutionary perspective. *Review of General Psychology*, *8*(2), 100–110. https://doi.org/10.1037/1089-2680.8.2.100

Dunham, Y. (2018). Mere membership. *Trends in Cognitive Sciences*, *22*(9), 780–793. https://doi.org/10.1016/j.tics.2018.06.004

Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliations in children. *Child Development*, *82*(3), 793–811. https://doi.org/10.1111/j.1467-8624.2011.01577.x

Eckstein, M. K., Guerra-Carrillo, B., Miller Singley, A. T., & Bunge, S. A. (2017). Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? *Developmental Cognitive Neuroscience*, *25*, 69–91. https://doi.org/10.1016/j.dcn.2016.11.001

Eggleston, A., Geangu, E., Tipper, S. P., Cook, R., & Over, H. (2021). Young children learn first impressions of faces through social referencing. *Scientific Reports*, *11*(1), 14744. https://doi.org/10.1038/s41598-021-94204-6

Eiser, J. R., & Stroebe, W. (1972). *Categorization and Social Judgment*. Academic Press.

Elbro, C., & Buch-Iversen, I. (2013). Activation of background knowledge for inference making: Effects on reading comprehension. *Scientific Studies of Reading*, *17*(6), 435–452. https://doi.org/10.1080/10888438.2013.774005

Esaulova, Y., Reali, C., & von Stockhausen, L. (2014). Influences of grammatical and stereotypical gender during reading: Eye movements in pronominal and noun phrase anaphor resolution. *Language, Cognition and Neuroscience*, *29*(7), 781–803. https://doi.org/10.1080/01690965.2013.794295

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., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and uses. *Annual Review of Psychology*, *54*(1), 297–327. https://doi.org/10.1146/annurev.psych.54.101601.145225

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

Fein, S., & Spencer, S. J. (1997). Prejudice as self-image maintenance: Affirming the self through derogating others. *Journal of Personality and Social Psychology*, *73*(1), 31–44. https://doi.org/10.1037/0022-3514.73.1.31

Ferera, M., Pun, A., Baron, A. S., & Diesendruck, G. (2021). The effect of familiarity on infants' social categorization capacity. *PLOS ONE*, *16*(3), e0247710. https://doi.org/10.1371/journal.pone.0247710

Filik, R. (2008). Contextual override of pragmatic anomalies: Evidence from eye movements. *Cognition*, *106*(2), 1038–1046. https://doi.org/10.1016/j.cognition.2007.04.006

Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In Gilbert. D. T., S. T. Fiske, & G. Lindzey (Eds.), *The Handbook of Social Psychology* (pp. 357–411). McGraw-Hill.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

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

Ford, T. E., & Stangor, C. (1992a). The role of diagnosticity in stereotype formation: Perceiving group means and variances. *Journal of Personality and Social Psychology*, *63*(3), 356–367. https://doi.org/10.1037/0022-3514.63.3.356

Ford, T. E., & Stangor, C. (1992b). The role of diagnosticity in stereotype formation: Perceiving group means and variances. *Journal of Personality and Social Psychology*, *63*(3), 356–367. https://doi.org/10.1037/0022-3514.63.3.356

Foster-Hanson, E., Leslie, S.-J., & Rhodes, M. (2019). Speaking of kinds: How generic language shapes the development of category representations. *Https://Doi.Org/10.31234/Osf.Io/28qf7*.

Foster-Hanson, E., & Rhodes, M. (2020). How origin stories shape children's social reasoning. *Cognitive Development*, *56*, 100962. https://doi.org/10.1016/j.cogdev.2020.100962

Foster-Hanson, E., & Rhodes, M. (2022). Stereotypes as prototypes in children's gender concepts. *Developmental Science*. https://doi.org/10.1111/desc.13345

Foster-Hanson, E., Roberts, S. O., Gelman, S. A., & Rhodes, M. (2021). Categories convey prescriptive information across domains and development. *Journal of Experimental Child Psychology*, *212*, 105231. https://doi.org/10.1016/j.jecp.2021.105231

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*(2), 178–210. https://doi.org/10.1016/0010-0285(82)90008-1

Gabriel, U., Behne, D. M., & Gygax, P. M. (2017). Speech vs. reading comprehension: An explorative study of gender representations in Norwegian. *Journal of Cognitive Psychology*, *29*(7), 795–808. https://doi.org/10.1080/20445911.2017.1326923

Gaertner, S. L., & Dovidio, J. F. (2004). Categorization, Recategorization, and Intergroup Bias. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the nature of prejudice: Fifty years after Allport* (pp. 71–88). Blackwell Publishing Ltd. https://doi.org/10.1002/9780470773963.ch5

Gaertner, S. L., & McLaughlin, J. P. (1983). Racial Stereotypes: Associations and Ascriptions of Positive and Negative Characteristics. *Social Psychology Quarterly*, *46*(1), 23. https://doi.org/10.2307/3033657

Garnham, A., Oakhill, J., & Reynolds, D. (2002). Are inferences from stereotyped role names to characters' gender made elaboratively? *Memory & Cognition*, *30*(3), 439–446. https://doi.org/10.3758/BF03194944

Garrod, S., & Sanford, A. (1977). Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behavior*, *16*(1), 77–90. https://doi.org/10.1016/S0022-5371(77)80009-1

Garrod, S., & Sanford, A. J. (1995). Incrementality in discourse understanding. In D. Milward & P. Sturt (Eds.), *Incremental interpretation: Edinburgh working papers in cognitive science* (pp. 99–122). Centre for Cognitive Science, The University of Edinburgh.

Gawronski, B., De Houwer, J., & Sherman, J. W. (2020). Twenty-five years of research using implicit measures. *Social Cognition*, *38*, S1–S25.

Gawronski, B., Hofmann, W., & Wilbur, C. J. (2006). Are "implicit" attitudes unconscious? *Consciousness and Cognition*, *15*(3), 485–499. https://doi.org/10.1016/j.concog.2005.11.007

Gawronski, B., & Payne, B. (2010). Handbook of Implicit Social Cognition: Measurement, theory, and applications. In B. Gawronski & B. K. Payne (Eds.), *Handbook of ImplIcIt SocIal CognItIon: Measurement, theory, and applications* . Guilford Press.

Gelman, S. A. (2003). *The Essential Child*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195154061.001.0001

Gelman, S. A., & Coley, J. D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology*, *26*(5), 796–804. https://doi.org/10.1037/0012-1649.26.5.796

Gelman, S. A., Collman, P., & Maccoby, E. E. (1986). Inferring properties from categories versus inferring categories from properties: The case of gender. *Child Development*, *57*(2), 396–404. https://doi.org/10.2307/1130595

Gelman, S. A., Goetz, P. J., Sarnecka, B. W., & Flukes, J. (2008). Generic language in parent-child conversations. *Language Learning and Development*, *4*(1), 1–31. https://doi.org/10.1080/15475440701542625

Gelman, S. A., & Heyman, G. D. (1999). Carrot-eaters and creature-believers: The effects of lexicalization on children's inferences about social categories. *Psychological Science*, *10*(6), 489–493. https://doi.org/10.1111/1467-9280.00194

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, *23*(3), 183–209. https://doi.org/10.1016/0010-0277(86)90034-X

Gelman, S. A., & Roberts, S. O. (2017). How language shapes the cultural inheritance of categories. *Proceedings of the National Academy of Sciences*, *114*(30), 7900–7907. https://doi.org/10.1073/pnas.1621073114

Gerrig, R. J., & McKoon, G. (1998). The readiness is all: The functionality of memory-based text processing. *Discourse Processes*, *26*(2–3), 67–86. https://doi.org/10.1080/01638539809545039

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

Glick, P., Zion, C., & Nelson, C. (1988). What mediates sex discrimination in hiring decisions? *Journal of Personality and Social Psychology*, *55*(2), 178–186. https://doi.org/10.1037/0022-3514.55.2.178

Gonzalez, A. M., Dunlop, W. L., & Baron, A. S. (2017). Malleability of implicit associations across development. *Developmental Science*, *20*(6), e12481. https://doi.org/10.1111/desc.12481

Gonzalez, C. M., Zosuls, K. M., & Ruble, D. N. (2010). Traits as dimensions or categories? Developmental change in the understanding of trait terms. *Developmental Psychology*, *46*(5), 1078–1088. https://doi.org/10.1037/a0020207

Gopnik, A. (2012). Scientific thinking in young children: Theoretical advances, empirical research, and policy implications. *Science*, *337*(6102), 1623–1627. https://doi.org/10.1126/science.1223416

Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology*, *37*(5), 620–629. https://doi.org/10.1037/0012-1649.37.5.620

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371–395. https://doi.org/10.1037/0033-295X.101.3.371

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(3), 501–518. https://doi.org/10.1037/0278-7393.11.3.501

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

Greenwald, A. G., & Banaji, M. R. (2017). The implicit revolution: Reconceiving the relation between conscious and unconscious. *American Psychologist*, *72*(9), 861–871. https://doi.org/10.1037/amp0000238

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

Greenwald, A. G., & Lai, C. K. (2020). Implicit social cognition. *Annual Review of Psychology*, *71*(1), 419–445. https://doi.org/10.1146/annurev-psych-010419-050837

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

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*(1), 17–41. https://doi.org/10.1037/a0015575

Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: Asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology*, *90*(1), 1–20. https://doi.org/10.1037/0022-3514.90.1.1

Grice, H. P. (1957). Meaning. *The Philosophical Review*, 377–388.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Speech acts* (Vol. 3, pp. 41–58). Academic Press.

Gygax, P., Sato, S., Öttl, A., & Gabriel, U. (2021). The masculine form in grammatically gendered languages and its multiple interpretations: A challenge for our cognitive system. *Language Sciences*, *83*, 101328. https://doi.org/10.1016/j.langsci.2020.101328

Hamilton, D. (1981). *Cognitive Processes in Stereotyping and Intergroup Behavior* (D. L. Hamilton, Ed.). Erlbaum. https://doi.org/10.4324/9781315668758

Hamilton, D. L., Dugan, P. M., & Trolier, T. K. (1985). The formation of stereotypic beliefs: Further evidence for distinctiveness-based illusory correlations. *Journal of Personality and Social Psychology*, *48*(1), 5–17. https://doi.org/10.1037/0022-3514.48.1.5

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology*, *12*(4), 392–407. https://doi.org/10.1016/S0022-1031(76)80006-6

Hamilton, D. L., & Rose, T. L. (1980). Illusory correlation and the maintenance of stereotypic beliefs. *Journal of Personality and Social Psychology*, *39*(5), 832–845. https://doi.org/10.1037/0022-3514.39.5.832

Harris, P. L. (2012). *Trusting what you're told: How children learn from others*. Harvard University Press. https://doi.org/10.4159/harvard.9780674065192

Haslam, N., Rothschild, L., & Ernst, D. (2002). Are essentialist beliefs associated with prejudice? *British Journal of Social Psychology*, *41*(1), 87–100. https://doi.org/10.1348/014466602165072

Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, *13*(5), 512–521.

Heim, I. (1982). *The semantics of definite and indefinite noun phrases* [Dissertation, University of Massachusetts]. Doctoral Dissertations Available from Proquest. AAI8229562.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Heyman, G. D., & Gelman, S. A. (1999). The use of trait labels in making psychological inferences. *Child Development*, *70*(3), 604–619. https://doi.org/10.1111/1467-8624.00044

Heyman, G. D., & Gelman, S. A. (2000a). Preschool children's use of novel predicates to make inductive inferences about people. *Cognitive Development*, *15*(3), 263–280. https://doi.org/10.1016/S0885-2014(00)00028-9

Heyman, G. D., & Gelman, S. A. (2000b). Preschool children's use of trait labels to make inductive inferences. *Journal of Experimental Child Psychology*, *77*(1), 1–19. https://doi.org/10.1006/jecp.1999.2555

Hilton, J. L., & von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology*, *47*(1), 237–271. https://doi.org/10.1146/annurev.psych.47.1.237

Hirschfeld, L. A. (1996). *Race in the making: Cognition, culture, and the child's construction of human kinds*. The MIT Press.

Hoicka, E., Saul, J., Prouten, E., Whitehead, L., & Sterken, R. (2021). Language signaling high proportions and generics lead to generalizing, but not essentializing, for novel social kinds. *Cognitive Science*, *45*(11), e13051. https://doi.org/10.1111/cogs.13051

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of inference, learning, and discovery.* . MIT Press.

Howard, J. W., & Rothbart, M. (1980). Social categorization and memory for in-group and out-group behavior. *Journal of Personality and Social Psychology*, *38*(2), 301–310. https://doi.org/10.1037/0022-3514.38.2.301

Inhoff, A. W., & Radach, R. (1998). Definition and computation of oculomotor measures in the study of cognitive processes. In *Eye Guidance in Reading and Scene Perception* (pp. 29–53). Elsevier. https://doi.org/10.1016/B978-008043361-5/50003-1

Irmen, L. (2007). What's in a (role) name? Formal and conceptual aspects of comprehending personal nouns. *Journal of Psycholinguistic Research*, *36*(6), 431–456. https://doi.org/10.1007/s10936-007-9053-z

Jacobs, J. E., & Narloch, R. H. (2001). Children's use of sample size and variability to make social inferences. *Journal of Applied Developmental Psychology*, *22*(3), 311–331. https://doi.org/10.1016/S0193-3973(01)00086-7

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., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*(3), 306–340. https://doi.org/10.1037/0096-3445.110.3.306

Jacoby, L. L., & Witherspoon, D. (1982). Remembering without awareness. *Canadian Journal of Psychology / Revue Canadienne de Psychologie*, *36*(2), 300–324. https://doi.org/10.1037/h0080638

Jaswal, V. K., & Markman, E. M. (2007). Looks aren't everything: 24-month-olds' willingness to accept unexpected labels. *Journal of Cognition and Development*, *8*(1), 93–111. https://doi.org/10.1080/15248370709336995

Jin, K., & Baillargeon, R. (2017). Infants possess an abstract expectation of ingroup support. *Proceedings of the National Academy of Sciences*, *114*(31), 8199–8204. https://doi.org/10.1073/pnas.1706286114

Johansen, M. K., & Kruschke, J. K. (2005). Category Representation for Classification and Feature Inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1433–1458. https://doi.org/10.1037/0278-7393.31.6.1433

Jordan, A., & Dunham, Y. (2021). Are category labels primary? Children use similarities to reason about social groups. *Developmental Science*, *24*(2), e13013. https://doi.org/10.1111/desc.13013

Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology*, *33*(1), 1–27. https://doi.org/10.1111/j.2044-8309.1994.tb01008.x

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A Decade of System Justification Theory: Accumulated Evidence of Conscious and Unconscious Bolstering of the Status Quo. *Political Psychology*, *25*(6), 881–919. https://doi.org/10.1111/j.1467-9221.2004.00402.x

Jost, J. T., & Burgess, D. (2000). Attitudinal ambivalence and the conflict between group and system justification motives in low status groups. *Personality and Social Psychology Bulletin*, *26*(3), 293–305. https://doi.org/10.1177/0146167200265003

Jost, J. T., & Hamilton, D. L. (2004). Stereotypes in Our Culture. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the nature of prejudice: Fifty years after Allport* (pp. 208–224). Blackwell Publishing Ltd. https://doi.org/10.1002/9780470773963.ch13

Jost, J. T., & Kay, A. C. (2005). Exposure to Benevolent Sexism and Complementary Gender Stereotypes: Consequences for Specific and Diffuse Forms of System Justification. *Journal of Personality and Social Psychology*, *88*(3), 498–509. https://doi.org/10.1037/0022-3514.88.3.498

Jost, J. T., Pelham, B. W., & Carvallo, M. R. (2002). Non-conscious forms of system justification: Implicit and behavioral preferences for higher status groups. *Journal of Experimental Social Psychology*, *38*(6), 586–602. https://doi.org/10.1016/S0022-1031(02)00505-X

Judd, C. M., & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review*, *100*(1), 109–128. https://doi.org/10.1037/0033-295X.100.1.109

Judd, C. M., & Park, B. (2004). Group differences and stereotype accuracy. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the nature of prejudice: Fifty years after Allport* (pp. 121–138). Blackwell Publishing Ltd. https://doi.org/10.1002/9780470773963.ch8

Juhasz, B. J., & Pollastek, A. (2011). Lexical influences on eye movements in reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 873–893). Oxford University Press.

Kalish, C. W. (2012). Generalizing norms and preferences within social categories and individuals. *Developmental Psychology*, *48*(4), 1133–1143. https://doi.org/10.1037/a0026344

Kalish, C. W., & Lawson, C. A. (2008). Development of social category representations: Early appreciation of roles and deontic relations. *Child Development*, *79*(3), 577–593. https://doi.org/10.1111/j.1467-8624.2008.01144.x

Kaufmann, L., & Clément, F. (2014). Wired for society: Cognizing pathways to society and culture. *Topoi*, *33*(2), 459–475. https://doi.org/10.1007/s11245-014-9236-9

Kay, A. C., & Jost, J. T. (2003). Complementary Justice: Effects of "Poor but Happy" and "Poor but Honest" Stereotype Exemplars on System Justification and Implicit Activation of the Justice Motive. *Journal of Personality and Social Psychology*, *85*(5), 823–837. https://doi.org/10.1037/0022-3514.85.5.823

Kay, A. C., Jost, J. T., & Young, S. (2005). Victim Derogation and Victim Enhancement as Alternate Routes to System Justification. *Psychological Science*, *16*(3), 240–246. https://doi.org/10.1111/j.0956-7976.2005.00810.x

Kendeou, P., McMaster, K. L., & Christ, T. J. (2016). Reading comprehension: Core components and processes. *Policy Insights from the Behavioral and Brain Sciences*, *3*(1), 62–69. https://doi.org/10.1177/2372732215624707

Kennison, S. M., & Trofe, J. L. (2003). Comprehending pronouns: A role for word-specific gender stereotype information. *Journal of Psycholinguistic Research*, *32*(3), 355–378. https://doi.org/10.1023/A:1023599719948

Kidder, C. K., White, K. R., Hinojos, M. R., Sandoval, M., & Crites, S. L. (2018). Sequential stereotype priming: A meta-analysis. *Personality and Social Psychology Review*, *22*(3), 199–227. https://doi.org/10.1177/1088868317723532

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences*, *104*(30), 12577–12580. https://doi.org/10.1073/pnas.0705345104

Kinzler, K. D., Shutts, K., & Correll, J. (2010). Priorities in social categories. *European Journal of Social Psychology*, *40*(4), 581–592. https://doi.org/10.1002/ejsp.739

Klein, O., & Bernard, P. (2015). Stereotypes in social psychology. In J. D. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 446–452). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.24097-X

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams, R. B., Alper, S., Aveyard, M., Axt, J. R., Babalola, M. T., Bahník, Š., Batra, R., Berkics, M., Bernstein, M. J., Berry, D. R., Bialobrzeska, O., Binan, E. D., Bocian, K., Brandt, M. J., Busching, R., … Nosek, B. A. (2018). Many Labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science*, *1*(4), 443–490. https://doi.org/10.1177/2515245918810225

Kliegl, R., Dambacher, M., Dimigen, O., Jacobs, A. M., & Sommer, W. (2012). Eye movements and brain electric potentials during reading. *Psychological Research*, *76*(2), 145–158. https://doi.org/10.1007/s00426-011-0376-x

Kramer, H. J., Goldfarb, D., Tashjian, S. M., & Hansen Lagattuta, K. (2021). Dichotomous thinking about social groups: Learning about one group can activate opposite beliefs about another group. *Cognitive Psychology*, *129*, 101408. https://doi.org/10.1016/j.cogpsych.2021.101408

Krieglmeyer, R., & Sherman, J. W. (2012). Disentangling stereotype activation and stereotype application in the stereotype misperception task. *Journal of Personality and Social Psychology*, *103*(2), 205–224. https://doi.org/10.1037/a0028764

Kurdi, B., & Banaji, M. (2022). Implicit social cognition: A brief (and gentle) introduction. In A. S. Reber & R. Allen (Eds.), *The cognitive unconscious: The first half-century* (pp. 323–346). Oxford University Press.

Kurdi, B., Mann, T. C., Charlesworth, T. E. S., & Banaji, M. R. (2019). The relationship between implicit intergroup attitudes and beliefs. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(13), 5862–5871. https://doi.org/10.1073/pnas.1820240116

Kurdi, B., Seitchik, A. E., Axt, J. R., Carroll, T. J., Karapetyan, A., Kaushik, N., Tomezsko, D., Greenwald, A. G., & Banaji, M. R. (2019). Relationship between the Implicit Association Test and intergroup behavior: A meta-analysis. *American Psychologist*, *74*(5), 569–586. https://doi.org/10.1037/amp0000364

Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science*, *21*(8), 1134–1140. https://doi.org/10.1177/0956797610376652

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470. https://doi.org/10.1016/S1364-6613(00)01560-6

Lane, J. D., Conder, E. B., & Rottman, J. (2020). The influence of direct and overheard messages on children's attitudes toward novel social groups. *Child Development*, *91*(3), 829–845. https://doi.org/10.1111/cdev.13238

Lassonde, K. A. (2015). Reducing the impact of stereotypical knowledge during reading. *Discourse Processes*, *52*(2), 149–171. https://doi.org/10.1080/0163853X.2014.917221

Lee, K., Quinn, P. C., & Pascalis, O. (2017). Face Race Processing and Racial Bias in Early Development: A Perceptual-Social Linkage. *Current Directions in Psychological Science*, *26*(3), 256–262. https://doi.org/10.1177/0963721417690276

Leinbach, M. D., Hort, B. E., & Fagot, B. I. (1997). Bears are for boys: Metaphorical associations in young children's gender stereotypes. *Cognitive Development*, *12*(1), 107–130. https://doi.org/10.1016/S0885-2014(97)90032-0

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

Leshin, R. A., Leslie, S., & Rhodes, M. (2021). Does it matter how we speak about social kinds? A large, preregistered, online experimental study of how language shapes the development of essentialist beliefs. *Child Development*, *92*(4), e531–e547. https://doi.org/10.1111/cdev.13527

Levy, S. R., Stroessner, S. J., & Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology*, *74*(6), 1421–1436. https://doi.org/10.1037/0022-3514.74.6.1421

Liberman, Z., Woodward, A. L., & Kinzler, K. D. (2017). The origins of social categorization. In *Trends in Cognitive Sciences* (Vol. 21, Issue 7, pp. 556–568). Elsevier Ltd. https://doi.org/10.1016/j.tics.2017.04.004

Lippmann, W. (1922). *Public Opinion*. Harcourt, Brace & Co.

Liu, D., Gelman, S. A., & Wellman, H. M. (2007). Components of young children's trait understanding: Behavior-to-trait inferences and trait-to-behavior predictions. *Child Development*, *78*(5), 1543–1558. https://doi.org/10.1111/j.1467-8624.2007.01082.x

Liversedge, S. P., Paterson, K. B., & Pickering, M. J. (1998). Eye movements and measures of reading time. In *Eye guidance in reading and scene perception* (pp. 55–75). Elsevier.

Loersch, C., & Keith Payne, B. (2016). Demystifying priming. *Current Opinion in Psychology*, *12*, 32–36. https://doi.org/10.1016/j.copsyc.2016.04.020

Ma, L., & Xu, F. (2011). Young children's use of statistical sampling evidence to infer the subjectivity of preferences. *Cognition*, *120*(3), 403–411. https://doi.org/10.1016/j.cognition.2011.02.003

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

Mahalingam, R. (2003). Essentialism, Culture, and Power: Representations of Social Class. *Journal of Social Issues*, *59*(4), 733–749. https://doi.org/10.1046/j.0022-4537.2003.00087.x

Mandalaywala, T. M., Amodio, D. M., & Rhodes, M. (2018). Essentialism Promotes Racial Prejudice by Increasing Endorsement of Social Hierarchies. *Social Psychological and Personality Science*, *9*(4), 461–469. https://doi.org/10.1177/1948550617707020

Mari, M. A. (2022). How cues to social categorization impact children's inferences about social categories. *Acta Psychologica*, *229*(103707), 1–12. https://doi.org/https://doi.org/10.1016/j.actpsy.2022.103707

Mari, M. A., & Müller, M.-L. (2023). Social cognition and Relevance: How stereotypes impact the processing of definite and indefinite descriptions. *Frontiers in Communication*, *8*, 1–16. https://doi.org/https://doi.org/10.3389/fcomm.2023.1088861

Markman, E. M. (1987). How children constrain the possible meanings of words. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 255–287). Cambridge University Press.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction.* MIT Press.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121–157. https://doi.org/10.1016/0010-0285(88)90017-5

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: facilitation and feature generalization. *Developmental Science*, *11*(1), 122–134. https://doi.org/10.1111/j.1467-7687.2007.00653.x

McGarty, C. (2002). Stereotype formation as category formation. In C. McGarty, V. Y. Yzerbyt, & R. Spears (Eds.), *Stereotypes as explanations: The formation of meaningful beliefs about social groups* (pp. 16–37). Cambridge University Press. https://doi.org/10.1017/CBO9780511489877.003

McGarty, C., Yzerbyt, V., & Spears, R. (2002). Stereotypes as explanations: The formation of meaningful beliefs about social groups. In C. McGarty, V. Y. Yzerbyt, & R. Spears (Eds.), *Stereotypes as explanations: The formation of meaningful beliefs about social groups.* Cambridge University Press. https://doi.org/10.1017/CBO9780511489877

Medin, D. L., Altom, M. W., & Murphy, T. D. (1984). Given versus induced category representations: Use of prototype and exemplar information in classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(3), 333–352. https://doi.org/10.1037/0278-7393.10.3.333

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*(3), 207–238. https://doi.org/10.1037/0033-295X.85.3.207

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, *34*(2), 57–74. https://doi.org/10.1017/S0140525X10000968

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227–234. https://doi.org/10.1037/h0031564

Misersky, J., Gygax, P. M., Canal, P., Gabriel, U., Garnham, A., Braun, F., Chiarini, T., Englund, K., Hanulikova, A., Öttl, A., Valdrova, J., von Stockhausen, L., & Sczesny, S. (2014). Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak. *Behavior Research Methods*, *46*(3), 841–871. https://doi.org/10.3758/s13428-013-0409-z

Morton, T. A., & Postmes, T. (2009). When Differences Become Essential: Minority Essentialism in Response to Majority Treatment. *Personality and Social Psychology Bulletin*, *35*(5), 656–668. https://doi.org/10.1177/0146167208331254

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

Moty, K., & Rhodes, M. (2021). The unintended consequences of the things we say: What generic statements communicate to children about unmentioned categories. *Psychological Science*, *32*(2), 189–203. https://doi.org/10.1177/0956797620953132

Müller, F., & Rothermund, K. (2014). What does it take to activate stereotypes? Simple primes don't seem enough. *Social Psychology*, *45*(3), 187–193. https://doi.org/10.1027/1864-9335/a000183

Müller, M.-L., & Mari, M. A. (2021). Definite descriptions in the light of the comprehension vs. acceptance distinction: Comparing self-paced reading with eye-tracking measures. *Frontiers in Communication*, *6*(634362), 1–17. https://doi.org/10.3389/fcomm.2021.634362

Murphy, G. L. (2016). Is there an exemplar theory of concepts? *Psychonomic Bulletin & Review*, *23*(4), 1035–1042. https://doi.org/10.3758/s13423-015-0834-3

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*(3), 289–316. https://doi.org/10.1037/0033-295X.92.3.289

Neely, J. H. (1976). Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. *Memory & Cognition*, *4*(5), 648–654. https://doi.org/10.3758/BF03213230

Newheiser, A.-K., Dunham, Y., Merrill, A., Hoosain, L., & Olson, K. R. (2014). Preference for high status predicts implicit outgroup bias among children from low-status groups. *Developmental Psychology*, *50*(4), 1081–1090. https://doi.org/10.1037/a0035054

Ni, W., Fodor, J. D., Crain, S., & Shankweiler, D. (1998). Anomaly detection: Eye movement patterns. *Journal of Psycholinguistic Research*, *27*(5), 515–539. https://doi.org/10.1023/A:1024996828734

Nosek, B. A., & Banaji, M. R. (2001). The Go/No-Go Association Task. *Social Cognition*, *19*(6), 625–666. https://doi.org/10.1521/soco.19.6.625.20886

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: From measures to mechanisms. *Trends in Cognitive Sciences*, *15*(4), 152–159. https://doi.org/10.1016/j.tics.2011.01.005

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., Smith, C. T., Olson, K. R., Chugh, D., Greenwald, A. G., & Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology*, *18*(1), 36–88. https://doi.org/10.1080/10463280701489053

Noyes, A., Dunham, Y., Keil, F. C., & Ritchie, K. (2021). Evidence for multiple sources of inductive potential: Occupations and their relations to social institutions. *Cognitive Psychology*, *130*, 101422. https://doi.org/10.1016/j.cogpsych.2021.101422

Noyes, A., & Keil, F. C. (2020). There is no privileged link between kinds and essences early in development. *Proceedings of the National Academy of Sciences*, *117*(20), 10633–10635. https://doi.org/10.1073/pnas.2003627117

Oakes, L. M., Horst, J. S., Kovack-Lesh, K. A., & Perone, S. (2009). How infants learn categories. In A. Woodward & A. Needham (Eds.), *Learning and the Infant Mind* (pp. 144–171). Oxford University Press.

O'Brien, E. J., Rizzella, M. L., Albrecht, J. E., & Halleran, J. G. (1998). Updating a situation model: A memory-based text processing view. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(5), 1200–1210. https://doi.org/10.1037/0278-7393.24.5.1200

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251). https://doi.org/10.1126/science.aac4716

Osterhout, L., Bersick, M., & Mclaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory & Cognition*, *25*(3), 273–285. https://doi.org/10.3758/BF03211283

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, *105*(2), 171–192. https://doi.org/10.1037/a0032734

Otten, S., & Wentura, D. (1999). About the impact of automaticity in the minimal group paradigm: evidence from affective priming tasks. *European Journal of Social Psychology*, *29*(8), 1049–1071. https://doi.org/10.1002/(SICI)1099-0992(199912)29:8<1049::AID-EJSP985>3.0.CO;2-Q

Over, H., & McCall, C. (2018). Becoming us and them: Social learning and intergroup bias. *Social and Personality Psychology Compass*, *12*(4), e12384. https://doi.org/10.1111/spc3.12384

Park, B., & Hastie, R. (1987). Perception of variability in category development: Instance-versus abstraction-based stereotypes. *Journal of Personality and Social Psychology*, *53*(4), 621–635. https://doi.org/10.1037/0022-3514.53.4.621

Park, B., & Judd, C. M. (2005). Rethinking the link between categorization and prejudice within the social cognition perspective. *Personality and Social Psychology Review*, *9*(2), 108–130. https://doi.org/10.1207/s15327957pspr0902_2

Pauker, K., Tai, C., & Ansari, S. (2020). *Contextualizing the development of social essentialism* (pp. 65–94). https://doi.org/10.1016/bs.acdb.2020.05.003

Paulus, M., Hunnius, S., van Wijngaarden, C., Vrins, S., van Rooij, I., & Bekkering, H. (2011). The role of frequency information and teleological reasoning in infants' and adults' action prediction. *Developmental Psychology*, *47*(4), 976–983. https://doi.org/10.1037/a0023785

Payne, B. K., & Gawronski, B. (2010). A history of implicit social cognition: Where is it coming from? Where is it now? Where is it going? In B. Gawronski & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, theory, and applications* (pp. 1–15). Guilford Press.

Peabody, D. (1967). Trait inferences: Evaluative and descriptive aspects. *Journal of Personality and Social Psychology*, *7*(4, Pt.2), 1–18. https://doi.org/10.1037/h0025230

Peretz-Lange, R., & Muentener, P. (2020). Children's use of generic labels, discreteness, and stability to form a novel category. *Journal of Cognition and Development*, *21*(3), 447–475. https://doi.org/10.1080/15248372.2020.1757452

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22–37. https://doi.org/10.1080/10888438.2013.827687

Pesciarelli, F., Scorolli, C., & Cacciari, C. (2019). Neural correlates of the implicit processing of grammatical and stereotypical gender violations: A masked and unmasked priming study. *Biological Psychology*, *146*, 107714. https://doi.org/10.1016/j.biopsycho.2019.06.002

Plötner, M., Over, H., Carpenter, M., & Tomasello, M. (2015). The effects of collaboration and minimal-group membership on children's prosocial behavior, liking, affiliation, and trust. *Journal of Experimental Child Psychology*, *139*, 161–173. https://doi.org/10.1016/j.jecp.2015.05.008

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium.* Erlbaum.

Powell, L. J., & Spelke, E. S. (2013). Preverbal infants expect members of social groups to act alike. *Proceedings of the National Academy of Sciences*, *110*(41). https://doi.org/10.1073/pnas.1304326110

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, *67*(4), 741–763. https://doi.org/10.1037/0022-3514.67.4.741

Proverbio, A. M., Alberio, A., & de Benedetto, F. (2018). Neural correlates of automatic beliefs about gender stereotypes: Males are more prejudicial. *Brain and Language*, *186*, 8–16. https://doi.org/10.1016/j.bandl.2018.08.006

Quadflieg, S., & Macrae, C. N. (2011). Stereotypes and stereotyping: What's the brain got to do with it? *European Review of Social Psychology*, *22*(1), 215–273. https://doi.org/10.1080/10463283.2011.627998

Quinn, P. C., Lee, K., Pascalis, O., & Tanaka, J. W. (2016). Narrowing in categorical responding to other-race face classes by infants. *Developmental Science*, *19*(3), 362–371. https://doi.org/10.1111/desc.12301

Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception*, *31*(9), 1109–1121. https://doi.org/10.1068/p3331

Rakoczy, H. (2022). Puppet studies present clear and distinct windows into the child's mind. *Cognitive Development*, *61*, 101147. https://doi.org/10.1016/j.cogdev.2021.101147

Rapp, D. N., & van den Broek, P. (2005). Dynamic text comprehension: An integrative view of eeading. *Current Directions in Psychological Science*, *14*(5), 276–279. https://doi.org/10.1111/j.0963-7214.2005.00380.x

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14*(3), 191–201. https://doi.org/10.3758/BF03197692

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision Research*, *16*(8), 829–837. https://doi.org/10.1016/0042-6989(76)90143-7

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(6), 1290–1301. https://doi.org/10.1037/0278-7393.30.6.1290

Reali, C., Esaulova, Y., Öttl, A., & von Stockhausen, L. (2015). Role descriptions induce gender mismatch effects in eye movements during reading. *Frontiers in Psychology*, *6*(1607), 1–13. https://doi.org/10.3389/fpsyg.2015.01607

Rees, H. R., Ma, D. S., & Sherman, J. W. (2020). Examining the relationships among categorization, stereotype activation, and stereotype application. *Personality and Social Psychology Bulletin*, *46*(4), 499–513. https://doi.org/10.1177/0146167219861431

Reichle, E. D. (2011). Serial-attention models of reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 767–786). Oxford University Press.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*(1), 125–157. https://doi.org/10.1037/0033-295X.105.1.125

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences*, *26*(4), 445–476. https://doi.org/10.1017/S0140525X03000104

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review*, *16*(1), 1–21. https://doi.org/10.3758/PBR.16.1.1

Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics*, *44*(6), 563–575. https://doi.org/10.3758/BF03207490

Reisenzein, R., Horstmann, G., & Schützwohl, A. (2019). The cognitive-evolutionary model of surprise: A review of the evidence. *Topics in Cognitive Science*, *11*(1), 50–74. https://doi.org/10.1111/tops.12292

Renkewitz, F., & Heene, M. (2019). The replication crisis and open science in psychology: Methodological challenges and developments. *Zeitschrift Für Psychologie*, *227*(4), 233–236. https://doi.org/10.1027/2151-2604/a000389

Reynolds, D., Garnham, A., & Oakhill, J. (2006). Evidence of immediate activation of gender information from a social role name. *Quarterly Journal of Experimental Psychology*, *59*(5), 886–903. https://doi.org/10.1080/02724980543000088

Rhodes, M. (2012). Naïve theories of social groups. *Child Development*, *83*(6), 1900–1916. https://doi.org/10.1111/j.1467-8624.2012.01835.x

Rhodes, M. (2013). How two intuitive theories shape the development of social categorization. *Child Development Perspectives*, *7*(1), 12–16. https://doi.org/10.1111/cdep.12007

Rhodes, M. (2014). Children's explanations as a window into their intuitive theories of the social world. *Cognitive Science*, *38*(8), 1687–1697. https://doi.org/10.1111/cogs.12129

Rhodes, M., & Baron, A. (2019). The development of social categorization. *Annual Review of Developmental Psychology*, *1*, 359–386. https://doi.org/10.1146/annurev-devpsych-121318- 084824

Rhodes, M., & Brickman, D. (2011). The influence of competition on children's social categories. *Journal of Cognition and Development*, *12*(2), 194–221. https://doi.org/10.1080/15248372.2010.535230

Rhodes, M., & Chalik, L. (2013). Social categories as markers of intrinsic interpersonal obligations. *Psychological Science*, *24*(6), 999–1006. https://doi.org/10.1177/0956797612466267

Rhodes, M., & Gelman, S. A. (2008). Categories influence predictions about individual consistency. *Child Development*, *79*(5), 1270–1287. https://doi.org/10.1111/j.1467-8624.2008.01188.x

Rhodes, M., Leslie, S. J., Bianchi, L., & Chalik, L. (2018). The role of generic language in the early development of social categorization. *Child Development*, *89*(1), 148–155. https://doi.org/10.1111/cdev.12714

Rhodes, M., Leslie, S., Saunders, K., Dunham, Y., & Cimpian, A. (2018). How does social essentialism affect the development of inter-group relations? *Developmental Science*, *21*(1). https://doi.org/10.1111/desc.12509

Rhodes, M., Leslie, S.-J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences*, *109*(34), 13526–13531. https://doi.org/10.1073/pnas.1208951109

Rhodes, M., Rizzo, M. T., Foster-Hanson, E., Moty, K., Leshin, R. A., Wang, M., Benitez, J., & Ocampo, J. D. (2020). Advancing developmental science via unmoderated remote research with children. *Journal of Cognition and Development*, *21*(4), 477–493. https://doi.org/10.1080/15248372.2020.1797751

Richter, N., Over, H., & Dunham, Y. (2016). The effects of minimal group membership on young preschoolers' social preferences, estimates of similarity, and behavioral attribution. *Collabra*, *2*(1). https://doi.org/10.1525/collabra.44

Riggs, A. E. (2019). Social statistics: Children use statistical reasoning to guide their inferences about the scope of social behavior. *Developmental Psychology*, *55*(1), 66–79. https://doi.org/10.1037/dev0000618

Riggs, A. E., Kalish, C. W., & Alibali, M. W. (2014a). Property content guides children's memory for social learning episodes. *Cognition*, *131*(2), 243–253. https://doi.org/10.1016/j.cognition.2014.01.004

Riggs, A. E., Kalish, C. W., & Alibali, M. W. (2014b). When you've seen one, have you seen them all? Children's memory for general and specific learning episodes. *Developmental Psychology*, *50*(6), 1653–1659. https://doi.org/10.1037/a0036130

Riggs, A. E., & Long, M. (2020). The Domain Frequency Association: A mental shortcut to guide children's generalization of norms and preferences. *Cognitive Development*, *54*, 100853. https://doi.org/10.1016/j.cogdev.2020.100853

Roberts, C. (2003). Uniqueness in Definite Noun Phrases. *Linguistics and Philosophy*, *26*(3), 287–350. https://doi.org/10.1023/A:1024157132393

Roberts, S. O., Gelman, S. A., & Ho, A. K. (2017). So it is, so it shall be: Group regularities license children's prescriptive judgments. *Cognitive Science*, *41*, 576–600. https://doi.org/10.1111/cogs.12443

Roberts, S. O., Ho, A. K., & Gelman, S. A. (2017). Group presence, category labels, and generic statements influence children to treat descriptive group regularities as prescriptive. *Journal of Experimental Child Psychology*, *158*, 19–31. https://doi.org/10.1016/j.jecp.2016.11.013

Robinson, H. M. (2005). *Unexpected (in)definiteness: Plural generic expressions in Romance*. The State University of New Jersey.

Rodríguez-Gómez, P., Romero-Ferreiro, V., Pozo, M. A., Hinojosa, J. A., & Moreno, E. M. (2020). Facing stereotypes: ERP responses to male and female faces after gender-stereotyped statements. *Social Cognitive and Affective Neuroscience*, *15*(9), 928–940. https://doi.org/10.1093/scan/nsaa117

Rosch, E. (1978). Principles of Categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Lawrence Erlbaum. https://doi.org/10.1016/B978-1-4832-1446-7.50028-5

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605. https://doi.org/10.1016/0010-0285(75)90024-9

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382–439. https://doi.org/10.1016/0010-0285(76)90013-X

Rosenberg, S., & Sedlak, A. (1972). *Structural representations of implicit personality theory* (pp. 235–297). https://doi.org/10.1016/S0065-2601(08)60029-5

Rothbart, M., & Taylor, M. (1992). Category labels and social reality: Do we view social categories as natural kinds? In G. R. Semin & K. Fiedler (Eds.), *Language, interaction and social cognition* (pp. 11–36). Sage Publications, Inc.

Rubio-Fernández, P. (2013). Associative and inferential processes in pragmatic enrichment: The case of emergent properties. *Language and Cognitive Processes*, *28*(6), 723–745. https://doi.org/10.1080/01690965.2012.659264

Ryazanov, A. A., & Christenfeld, N. J. S. (2018). The strategic value of essentialism. *Social and Personality Psychology Compass*, *12*(1), e12370. https://doi.org/10.1111/spc3.12370

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Salomon, E., & Cimpian, A. (2014). The inherence heuristic as a source of essentialist thought. *Personality and Social Psychology Bulletin*, *40*(10), 1297–1315. https://doi.org/10.1177/0146167214541659

Sanbonmatsu, D. M., Sherman, S. J., & Hamilton, D. L. (1987). Illusory Correlation in the Perception of Individuals and Groups. *Social Cognition*, *5*(1), 1–25. https://doi.org/10.1521/soco.1987.5.1.1

Sato, S., Gygax, P. M., & Gabriel, U. (2013). Gender inferences: Grammatical features and their impact on the representation of gender in bilinguals. *Bilingualism: Language and Cognition*, *16*(4), 792–807. https://doi.org/10.1017/S1366728912000739

Saucier, D. A., Miller, C. T., & Doucet, N. (2005). Differences in Helping Whites and Blacks: A Meta-Analysis. *Personality and Social Psychology Review*, *9*(1), 2–16. https://doi.org/10.1207/s15327957pspr0901_1

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(3), 501–518. https://doi.org/10.1037/0278-7393.13.3.501

Schacter, D. L., & Graf, P. (1986). Effects of elaborative processing on implicit and explicit memory for new associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(3), 432–444. https://doi.org/10.1037/0278-7393.12.3.432

Schaller, M., & Maass, A. (1989). Illusory correlation and social categorization: Toward an integration of motivational and cognitive factors in stereotype formation. *Journal of Personality and Social Psychology*, *56*(5), 709–721. https://doi.org/10.1037/0022-3514.56.5.709

Schidelko, L. P., Schünemann, B., Rakoczy, H., & Proft, M. (2021). Online testing yields the same results as lab testing: A validation study with the False Belief Task. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.703238

Schwarz, F. (2009). Two Types of Definites in Natural Language. *Open Access Dissertations*. https://scholarworks.umass.edu/open_access_dissertations/122

Schwarz, F. (2019). Definites, domain restriction, and discourse structure in online processing. In C. J. Clifton, K. Carlson, & J. D. Fodor (Eds.), *Grammatical Approaches to Language Processing* (pp. 187–208). Springer. https://doi.org/10.1007/978-3-030-01563-3_10

Segall, G., Birnbaum, D., Deeb, I., & Diesendruck, G. (2015). The intergenerational transmission of ethnic essentialism: How parents talk counts the most. *Developmental Science*, *18*(4), 543–555. https://doi.org/10.1111/desc.12235

Sherman, J. W., Klein, S. B., Laskey, A., & Wyer, N. A. (1998). Intergroup Bias in Group Judgment Processes: The Role of Behavioral Memories. *Journal of Experimental Social Psychology*, *34*(1), 51–65. https://doi.org/10.1006/jesp.1997.1342

Sherman, J. W., Kruschke, J. K., Sherman, S. J., Percy, E. J., Petrocelli, J. v., & Conrey, F. R. (2009). Attentional processes in stereotype formation: A common model for category accentuation and illusory correlation. *Journal of Personality and Social Psychology*, *96*(2), 305–323. https://doi.org/10.1037/a0013778

Sherman, J. W., Lee, A. Y., Bessenoff, G. R., & Frost, L. A. (1998). Stereotype efficiency reconsidered: Encoding flexibility under cognitive load. *Journal of Personality and Social Psychology*, *75*(3), 589–606. https://doi.org/10.1037/0022-3514.75.3.589

Sherman, S. J., Sherman, J. W., Percy, E. J., & Soderberg, C. K. (2013). Stereotype Development and Formation. In D. Carlston (Ed.), *The Oxford Handbook of Social Cognition* (pp. 548–574). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199730018.013.0027

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

Shilo, R., Weinsdörfer, A., Rakoczy, H., & Diesendruck, G. (2021). Children's prediction of others' behavior based on group vs. individual properties. *Cognitive Development*, *57*, 100955. https://doi.org/10.1016/j.cogdev.2020.100955

Shutts, K., Banaji, M. R., & Spelke, E. S. (2009). Social categories guide young children's preferences for novel objects. *Developmental Science*, *13*(4), 599–610. https://doi.org/10.1111/j.1467-7687.2009.00913.x

Shutts, K., & Kalish, C. W. (2021). Intuitive sociology. In *Advances in Child Development and Behavior* (Vol. 61, pp. 335–374). Academic Press Inc. https://doi.org/10.1016/bs.acdb.2021.05.004

Shutts, K., Roben, C. K. P., & Spelke, E. S. (2013). Children's use of social categories in thinking about people and social relationships. *Journal of Cognition and Development*, *14*(1), 35–62. https://doi.org/10.1080/15248372.2011.638686

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression.* Cambridge University Press. https://doi.org/10.1017/CBO9781139175043

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *Quarterly Journal of Economics*, *69*, 99–118.

Singh, R., Fedorenko, E., Mahowald, K., & Gibson, E. (2016). Accommodating presuppositions is inappropriate in implausible contexts. *Cognitive Science*, *40*(3), 607–634. https://doi.org/10.1111/cogs.12260

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2016). *afex: Analysis of Factorial Experiments*.

Skinner, A. L., & Meltzoff, A. N. (2019). Childhood experiences and intergroup biases among children. *Social Issues and Policy Review*, *13*(1), 211–240. https://doi.org/10.1111/sipr.12054

Skinner, A. L., Meltzoff, A. N., & Olson, K. R. (2017). "Catching" social bias: Exposure to biased nonverbal signals creates social biases in preschool children. *Psychological Science*, *28*(2), 216–224. https://doi.org/10.1177/0956797616678930

Skinner, A. L., Olson, K. R., & Meltzoff, A. N. (2020). Acquiring group bias: Observing other people's nonverbal signals can create social group biases. *Journal of Personality and Social Psychology*, *119*(4), 824–838. https://doi.org/10.1037/pspi0000218

Sloutsky, V. M., & Fisher, A. v. (2004). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology: General*, *133*(2), 166–188. https://doi.org/10.1037/0096-3445.133.2.166

Sloutsky, V. M., Kloos, H., & Fisher, A. v. (2007). When looks are everything. *Psychological Science*, *18*(2), 179–185. https://doi.org/10.1111/j.1467-9280.2007.01869.x

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Harvard University Press.

Smith, E. R., & Zárate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review*, *99*(1), 3–21. https://doi.org/10.1037/0033-295X.99.1.3

Sniderman, P. M., & Piazza, T. (1993). *The scar of race*. Belknap Press.

Soley, G. (2019). What do group members share? The privileged status of cultural knowledge for children. *Cognitive Science*, *43*(10). https://doi.org/10.1111/cogs.12786

Sparks, E., Schinkel, M. G., & Moore, C. (2017). Affiliation affects generosity in young children: The roles of minimal group membership and shared interests. *Journal of Experimental Child Psychology*, *159*, 242–262. https://doi.org/10.1016/j.jecp.2017.02.007

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. In *Developmental Science* (Vol. 10, Issue 1, pp. 89–96). https://doi.org/10.1111/j.1467-7687.2007.00569.x

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic Vigilance. *Mind and Language*, *25*(4), 359–393. https://doi.org/10.1111/j.1468-0017.2010.01394.x

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition* (Vol. 142). Harvard University Press.

Sperber, D., & Wilson, D. (2015). Beyond speaker's meaning. *Croatian Journal of Philosophy*, *15*(2 (44)), 117–149. https://hrcak.srce.hr/167180

Stanovich, K. E., West, R. F., & Toplak, M. E. (2014). Rationality, intelligence, and the defining features of Type 1 and Type 2 processing. In *Dual-process theories of the social mind.* (pp. 80–91). The Guilford Press.

Stroebe, W., & Insko, C. A. (1989). Stereotype, Prejudice, and Discrimination: Changing Conceptions in Theory and Research. In D. Bar-Tal, C. F. Graumann, A. W. Kruglanski, & W. Stroebe (Eds.), *Stereotyping and Prejudice* (pp. 3–34). Springer New York. https://doi.org/10.1007/978-1-4612-3582-8_1

Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Biosocial Science*, *1*(S1), 173–191. https://doi.org/10.1017/S0021932000023336

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, *223*(5), 96–103.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*(2), 149–178. https://doi.org/10.1002/ejsp.2420010202

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations*. Brooks/Cole.

Tajfel, H., & Turner, J. C. (1985). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *The psychology of intergroup relations* (pp. 7–24). Nelson-Hall.

Tajfel, H., & Wilkes, A. L. (1963). Classification and quantitative judgement. *British Journal of Psychology*, *54*(2), 101–114. https://doi.org/10.1111/j.2044-8295.1963.tb00865.x

Thalmayer, A. G., Toscanelli, C., & Arnett, J. J. (2021). The neglected 95% revisited: Is American psychology becoming less American? *American Psychologist*, *76*(1), 116–129. https://doi.org/10.1037/amp0000622

Vales, C., Wu, C., Torrance, J., Shannon, H., States, S. L., & Fisher, A. v. (2021). Research at a distance: Replicating semantic differentiation effects using remote data collection with children participants. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.697550

Van Bavel, J. J., & Cunningham, W. A. (2009). Self-Categorization With a Novel Mixed-Race Group Moderates Automatic Social and Racial Biases. *Personality and Social Psychology Bulletin*, *35*(3), 321–335. https://doi.org/10.1177/0146167208327743

van Knippenberg, A., Dijksterhuis, A., & Vermeulen, D. (1999). Judgement and memory of a criminal act: The effects of stereotypes and cognitive load. *European Journal of Social Psychology*, *29*(2–3), 191–201. https://doi.org/10.1002/(SICI)1099-0992(199903/05)29:2/3<191::AID-EJSP923>3.0.CO;2-O

Vasilyeva, N., Gopnik, A., & Lombrozo, T. (2018). The development of structural thinking about social categories. *Developmental Psychology*, *54*(9), 1735–1744. https://doi.org/10.1037/dev0000555

Vauras, M., Hyona, J., & Niemi, P. (1992). Comprehending coherent and incoherent texts: Evidence from eye movement patterns and recall performance. *Journal of Research in Reading*, *15*(1), 39–54. https://doi.org/10.1111/j.1467-9817.1992.tb00020.x

Vitu, F., O'Regan, J. K., Inhoff, A. W., & Topolski, R. (1995). Mindless reading: Eye-movement characteristics are similar in scanning letter strings and reading texts. *Perception & Psychophysics*, *57*(3), 352–364. https://doi.org/10.3758/BF03213060

Walton, G. M., & Banaji, M. R. (2004). Being what you say: The effect of essentialist linguistic labels on preferences. *Social Cognition*, *22*(2), 193–213. https://doi.org/10.1521/soco.22.2.193.35463

Warren, T. (2011). The influence of implausibility and anomaly on eye movements during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 911–923). Oxford University Press.

Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. *Psychonomic Bulletin & Review*, *14*(4), 770–775. https://doi.org/10.3758/BF03196835

Warren, T., McConnell, K., & Rayner, K. (2008). Effects of context on eye movements when reading about possible and impossible events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 1001–1010. https://doi.org/10.1037/0278-7393.34.4.1001

Waxman, S. R. (2010). Names will never hurt me? Naming and the development of racial and gender categories in preschool-aged children. *European Journal of Social Psychology*, *40*(4), 593–610. https://doi.org/10.1002/ejsp.732

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology*, *29*(3), 257–302. https://doi.org/10.1006/cogp.1995.1016

Weinraub, M., Clemens, L. P., Sockloff, A., Ethridge, T., Gracely, E., & Myers, B. (1984). The development of sex role stereotypes in the third year: Relationships to gender labeling, gender identity, sex-types toy preference, and family characteristics. *Child Development*, *55*(4), 1493–1503. https://doi.org/10.2307/1130019

Weisbuch, M., & Ambady, N. (2008). Non-conscious routes to building culture: Nonverbal components of socialization. *Journal of Consciousness Studies*, *15*(10–11), 159–183.

Wellman, H. M., Kushnir, T., Xu, F., & Brink, K. A. (2016). Infants use statistical sampling to understand the psychological world. *Infancy*, *21*(5), 668–676. https://doi.org/10.1111/infa.12131

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

Wharton, T., & Strey, C. (2019). Slave of the Passions: Making Emotions Relevant. In K. Scott, B. Clark, & R. Carston (Eds.), *Relevance, Pragmatics and Interpretation* (1st ed., pp. 253–266). Cambridge University Press. https://doi.org/10.1017/9781108290593.022

White, K. R., Crites, S. L., Taylor, J. H., & Corral, G. (2009). Wait, what? Assessing stereotype incongruities using the N400 ERP component. *Social Cognitive and Affective Neuroscience*, *4*(2), 191–198. https://doi.org/10.1093/scan/nsp004

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag.

Wilks, M., & Nielsen, M. (2018). Children disassociate from antisocial in-group members. *Journal of Experimental Child Psychology*, *165*, 37–50. https://doi.org/10.1016/j.jecp.2017.06.003

Wilson, D. (2016). Reassessing the conceptual–procedural distinction. *Lingua*, *175*, 5–19.

Winter, B. (2019). *Statistics for linguists: An introduction using R*. Routledge. https://doi.org/10.4324/9781315165547

Winter, B., & Wieling, M. (2016). How to analyze linguistic change using mixed models, Growth Curve Analysis and Generalized Additive Modeling. *Journal of Language Evolution*, *1*(1), 7–18. https://doi.org/10.1093/jole/lzv003

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, *72*(2), 262–274. https://doi.org/10.1037/0022-3514.72.2.262

Wyer, R. S., & Carlston, D. E. (2018). *Social Cognition, Inference, and Attribution*. Erlbaum. https://doi.org/10.4324/9780203781593

Yamamoto, H. W., Kawahara, M., & Tanaka, A. (2021). A web-based auditory and visual emotion perception task experiment with children and a comparison of lab data and web data. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.702106

Yamauchi, T., & Markman, A. B. (1998). Category Learning by Inference and Classification. *Journal of Memory and Language*, *39*(1), 124–148. https://doi.org/10.1006/jmla.1998.2566

Yzerbyt, V., Corneille, O., & Estrada, C. (2001). The Interplay of Subjective Essentialism and Entitativity in the Formation of Stereotypes. *Personality and Social Psychology Review*, *5*(2), 141–155. https://doi.org/10.1207/S15327957PSPR0502_5

Yzerbyt, V., Rocher, S., & Schadron, G. (1997). Stereotypes as explanations: A subjective essentialistic view of group perception. In R. Spears, P. J. Oakes, N. Ellemers, & S. A. Haslam (Eds.), *The social psychology of stereotyping and group life* (pp. 20–50). Blackwell.

Zosh, J. M., Brinster, M., & Halberda, J. (2013). Optimal contrast: Competition between two referents improves word learning. *Applied Developmental Science*, *17*(1), 20–28. https://doi.org/10.1080/10888691.2013.748420